

Rec'd PCT/PTC 17 FEB 2005  
ST/EP/03/09105 #2

# BUNDESREPUBLIK DEUTSCHLAND



REC'D - 3 OCT 2003	
WIPO	PCT

PCT/EP03/9105

## Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

**Aktenzeichen:** 102 53 112.9

**Anmeldetag:** 13. November 2002

**Anmelder/Inhaber:** SunGene GmbH & Co. KGaA, Gatersleben/DE

**Bezeichnung:** Verfahren zur Herstellung von Ketocarotinoiden in genetisch veränderten Organismen

**IPC:** C 12 N, C 12 P, A 01 H

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 10. September 2003  
Deutsches Patent- und Markenamt  
Der Präsident  
Im Auftrag

snypc0143

**PRIORITY DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH  
RULE 17.1(a) OR (b)

**BEST AVAILABLE COPY**

## Patentansprüche

1. Verfahren zur Herstellung von Ketocarotinoiden durch Kultivierung von genetisch veränderten Organismen, die im Vergleich zum Wildtyp eine veränderte Ketolase-Aktivität aufweisen, und die veränderte Ketolase-Aktivität durch eine Ketolase verursacht wird, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass man Organismen verwendet, die als Wildtyp bereits eine Ketolase-Aktivität aufweisen, und die genetische Veränderung eine Erhöhung der Ketolase-Aktivität im Vergleich zum Wildtyp bewirkt.
3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, dass man zur Erhöhung der Ketolase-Aktivität die Genexpression einer Nukleinsäure, kodierend eine Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist, gegenüber dem Wildtyp erhöht.
4. Verfahren nach Anspruch 3, dadurch gekennzeichnet, dass man zur Erhöhung der Genexpression Nukleinsäuren in den Organismus einbringt, die Ketolasen kodieren, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.
5. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass man Organismen verwendet, die als Wildtyp keine Ketolase-Aktivität aufweisen und die genetische Veränderung eine Ketolase-Aktivität im Vergleich zum Wildtyp verursacht.
6. Verfahren nach Anspruch 5, dadurch gekennzeichnet, dass man genetisch veränderte Organismen verwendet, die transgen eine Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist, exprimieren.
7. Verfahren nach Anspruch 5 oder 6, dadurch gekennzeichnet, dass man zur Verursachung der Genexpression Nukleinsäuren in die Organismen einbringt, die Ketolasen kodieren, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch

## 2

Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist

- 5 8. Verfahren nach Anspruch 5 oder 7, dadurch gekennzeichnet, dass man Nukleinsäuren, enthaltend die Sequenz SEQ. ID. NO. 1 einbringt.
- 10 9. Verfahren nach einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, dass die Organismen zusätzlich gegenüber dem Wildtyp eine erhöhte Aktivität mindestens einer der Aktivitäten, ausgewählt aus der Gruppe Hydroxylase-Aktivität und  $\beta$ -Cyclase-Aktivität, aufweisen.
- 15 10. Verfahren nach Anspruch 9, dadurch gekennzeichnet, dass man zur zusätzlichen Erhöhung mindestens einer der Aktivitäten, die Genexpression mindestens einer Nukleinsäure ausgewählt aus der Gruppe Nukleinsäuren, kodierend eine Hydroxylase, und Nukleinsäuren, kodierend eine  $\beta$ -Cyclase, gegenüber dem Wildtyp erhöht.
- 20 11. Verfahren nach Anspruch 10, dadurch gekennzeichnet, dass man zur Erhöhung der Genexpression mindestens eine Nukleinsäure ausgewählt aus der Gruppe, Nukleinsäuren kodierend eine Hydroxylase und Nukleinsäuren kodierend eine  $\beta$ -Cyclase in den Organismus einbringt.
- 25 12. Verfahren nach Anspruch 11, dadurch gekennzeichnet, dass man als Nukleinsäure, kodierend eine Hydroxylase, Nukleinsäuren einbringt, die eine Hydroxylase kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 16 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 % auf Aminosäureebene mit der Sequenz SEQ ID NO: 16 aufweist.
- 30 13. Verfahren nach Anspruch 12, dadurch gekennzeichnet, dass man Nukleinsäuren, enthaltend die Sequenz SEQ ID NO: 15 einbringt.
- 35 14. Verfahren nach Anspruch 11, dadurch gekennzeichnet, dass man als Nukleinsäure, kodierend eine  $\beta$ -Cyclase, Nukleinsäuren einbringt, die eine  $\beta$ -Cyclase kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 18 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 % auf Aminosäureebene mit der Sequenz SEQ ID NO: 18 aufweist.
15. Verfahren nach Anspruch 14, dadurch gekennzeichnet, dass man Nukleinsäuren, enthaltend die Sequenz SEQ ID NO: 17 einbringt.

## 3

16. Verfahren nach einem der Ansprüche 1 bis 15, dadurch gekennzeichnet, dass man nach dem Kultivieren die genetisch veränderten Organismen erntet und anschließend die Keto-carotinoide aus den Organismen isoliert.
- 5 17. Verfahren nach einem der Ansprüche 1 bis 16, dadurch gekennzeichnet daß man als Organismus einen Organismus verwendet, der als Ausgangsorganismus natürlicherweise oder durch genetische Komplementierung oder Umregulierung von Stoffwechselwegen in der Lage ist, Carotinoide herzustellen.
- 10 18. Verfahren nach einem der Ansprüche 1 bis 17, dadurch gekennzeichnet, daß man als Organismen Mikroorganismen oder Pflanzen verwendet.
- 15 19. Verfahren nach Anspruch 18, dadurch gekennzeichnet, daß man als Mikroorganismen Bakterien, Hefen, Algen oder Pilze verwendet.
- 20 20. Verfahren nach Anspruch 19, dadurch gekennzeichnet, daß die Mikroorganismen ausgewählt sind aus der Gruppe *Escherichia*, *Erwinia*, *Agrobacterium*, *Flavobacterium*, *Alcaligenes*, *Paracoccus*, *Nostoc*, Cyanobakterien der Gattung *Synechocystis*, *Candida*, *Saccharomyces*, *Hansenula*, *Phaffia*, *Pichia*, *Aspergillus*, *Trichoderma*, *Ashbya*, *Neurospora*, *Blakeslea*, *Phycomyces*, *Fusarium*, *Haematococcus*, *Phaedactylum tricornatum*, *Volvox* oder *Dunaliella*.
- 25 21. Verfahren nach Anspruch 18, dadurch gekennzeichnet, daß man als Organismus Pflanzen verwendet.
- 30 22. Verfahren nach Anspruch 21, dadurch gekennzeichnet, dass man als Pflanze eine Pflanze, ausgewählt aus den Familien Ranunculaceae, Berberidaceae, Papaveraceae, Cannabaceae, Rosaceae, Fabaceae, Linaceae, Vitaceae, Brassiceae, Cucurbitaceae, Primulaceae, Caryophyllaceae, Amaranthaceae, Gentianaceae, Geraniaceae, Caprifoliaceae, Oleaceae, Tropaeolaceae, Solanaceae, Scrophulariaceae, Asteraceae, Liliaceae, Amaryllidaceae, Poaceae, Orchidaceae, Malvaceae, Iliaceae oder Lamiaceae verwendet.
- 35 23. Verfahren nach Anspruch 22, dadurch gekennzeichnet, dass man als Pflanze eine Pflanze, ausgewählt aus den Pflanzengattungen Marigold, *Tagetes erecta*, *Tagetes patula*, *Acacia*, *Aconitum*, *Adonis*, *Arnica*, *Aquilegia*, *Aster*, *Astragalus*, *Bignonia*, *Calendula*, *Caltha*, *Campanula*, *Canna*, *Centaurea*, *Cheiranthus*, *Chrysanthemum*, *Citrus*, *Crepis*, *Crocus*, *Curcubita*, *Cytisus*, *Delonia*, *Delphinium*, *Dianthus*, *Dimorphotheca*, *Doronicum*, *Eschscholtzia*, *Forsythia*, *Fremontia*, *Gazania*, *Gelsemium*, *Genista*, *Gentiana*, *Geranium*, *Gerbera*, *Geum*, *Grevillea*, *Helenium*, *Helianthus*, *Hepatica*, *Heracleum*, *Hisbiscus*, *Heliopsis*, *Hypericum*, *Hypochoeris*, *Impatiens*, *Iris*, *Jacaranda*, *Kerria*, *Laburnum*, *Lathyrus*, *Leontodon*, *Lilium*, *Li-*

## 4

num, Lotus, Lycopersicon, Lysimachia, Maratia, Medicago, Mimulus, Narcissus, Oenothera, Osmanthus, Petunia, Photinia, Physalis, Phyteuma, Potentilla, Pyracantha, Ranunculus, Rhododendron, Rosa, Rudbeckia, Senecio, Silene, Silphium, Sinapsis, Sorbus, Spartium, Tecoma, Torenia, Tragopogon, Trollius, Tropaeolum, Tulipa, Tussilago, Ulex, Viola oder Zinnia verwendet.

24. Verfahren nach einem der Ansprüche 1 bis 23, dadurch gekennzeichnet, dass die Ketocarotinoide ausgewählt sind aus der Gruppe Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin und Adonixanthin.

25. Genetisch veränderter Organismus, wobei die genetische Veränderung die Aktivität einer Ketolase

A für den Fall, dass der Wildtyporganismus bereits eine Ketolase-Aktivität aufweist, gegenüber dem Wildtyp erhöht und

B für den Fall, dass der Wildtyporganismus keine Ketolase-Aktivität aufweist, gegenüber dem Wildtyp verursacht

und die nach A erhöhte oder nach B verursachte Ketolase-Aktivität durch eine Ketolase verursacht wird, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.

26. Genetisch veränderter Organismus nach Anspruch 25, dadurch gekennzeichnet, dass die Erhöhung oder Verursachung der Ketolase-Aktivität durch eine Erhöhung oder Verursachung der Genexpression einer Nukleinsäure, kodierend eine Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist, gegenüber dem Wildtyp bewirkt wird.

27. Genetisch veränderter Organismus nach Anspruch 26, dadurch gekennzeichnet, dass man zur Erhöhung oder Verursachung der Genexpression Nukleinsäuren in den Organismus einbringt, die Ketolasen kodieren, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der

Sequenz SEQ. ID. NO. 2 aufweist.

- 5 28. Genetisch veränderter Organismus, enthaltend mindestens eine transgene Nukleinsäure, kodierend eine Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.
- 10 29. Genetisch veränderter Organismus, enthaltend mindestens zwei endogene Nukleinsäuren, kodierend eine Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.
- 15 30. Genetisch veränderter Organismus nach einem der Ansprüche 25 bis 29, dadurch gekennzeichnet, dass die genetische Veränderung zusätzlich mindestens eine der Aktivitäten, ausgewählt aus der Gruppe Hydroxylase-Aktivität und  $\beta$ -Cyclase-Aktivität gegenüber dem Wildtyp erhöht.
- 20 31. Genetisch veränderter Organismus nach einem der Ansprüche 25 bis 30, dadurch gekennzeichnet daß er als Ausgangsorganismus natürlicherweise oder durch genetische Komplementation in der Lage ist, Carotinoide herzustellen.
- 25 32. Genetisch veränderter Organismus nach einem der Ansprüche 25 bis 31, ausgewählt aus der Gruppe Mikroorganismen oder Pflanzen.
- 30 33. Genetisch veränderter Mikroorganismus nach Anspruch 32, dadurch gekennzeichnet, daß die Mikroorganismen ausgewählt sind aus der Gruppe Bakterien, Hefen, Algen oder Pilze.
- 35 34. Genetisch veränderter Mikroorganismus nach Anspruch 33, dadurch gekennzeichnet, daß die Mikroorganismen ausgewählt sind aus der Gruppe *Escherichia*, *Erwinia*, *Agrobacterium*, *Flavobacterium*, *Alcaligenes*, *Paracoccus*, *Nostoc*, Cyanobakterien der Gattung *Synechocystis*, *Candida*, *Saccharomyces*, *Hansenula*, *Pichia*, *Aspergillus*, *Trichoderma*, *Ashbya*, *Neurospora*, *Blakeslea*, *Phycomyces*, *Fusarium*, *Haematococcus*, *Phaedactylum tricornatum*, *Volvox* oder *Dunaliella*.
- 35 35. Genetisch veränderte Pflanze nach Anspruch 32, dadurch gekennzeichnet, dass die Pflanzen ausgewählt sind aus den Familien Ranunculaceae, Berberidaceae, Papaveraceae, Cannabaceae, Rosaceae, Fabaceae, Linaceae, Vitaceae, Brassiceae, Cucurbitaceae, Pri-

mulaceae, Caryophyllaceae, Amaranthaceae, Gentianaceae, Geraniaceae, Caprifoliaceae, Oleaceae, Tropaeolaceae, Solanaceae, Scrophulariaceae, Asteraceae, Liliaceae, Amaryllidaceae, Poaceae, Orchidaceae, Malvaceae, Iliaceae oder Lamiaceae verwendet.

- 5 36. Genetisch veränderte Pflanze nach Anspruch 35, dadurch gekennzeichnet, dass Pflanzen ausgewählt sind aus den Pflanzengattungen Marigold, *Tagetes erecta*, *Tagetes patula*, *A-*  
10 *cacia*, *Aconitum*, *Adonis*, *Arnica*, *Aquilegia*, *Aster*, *Astragalus*, *Bignonia*, *Calendula*, *Caltha*,  
*Campanula*, *Canna*, *Centaurea*, *Cheiranthus*, *Chrysanthemum*, *Citrus*, *Crepis*, *Crocus*,  
15 *Curcubita*, *Cytisus*, *Delonia*, *Delphinium*, *Dianthus*, *Dimorphotheca*, *Doronicum*, *Esch-*  
*scholtzia*, *Forsythia*, *Fremontia*, *Gazania*, *Gelsemium*, *Genista*, *Gentiana*, *Geranium*, *Ger-*  
*bera*, *Geum*, *Grevillea*, *Helenium*, *Helianthus*, *Hepatica*, *Heracleum*, *Hisbiscus*, *Heliopsis*,  
*Hypericum*, *Hypochoeris*, *Impatiens*, *Iris*, *Jacaranda*, *Kerria*, *Laburnum*, *Lathyrus*, *Leonto-*  
20 *don*, *Lilium*, *Linum*, *Lotus*, *Lycopersicon*, *Lysimachia*, *Maratia*, *Medicago*, *Mimulus*, *Narcis-*  
*sus*, *Oenothera*, *Osmanthus*, *Petunia*, *Photinia*, *Physalis*, *Phyteuma*, *Potentilla*, *Pyracantha*,  
*Ranunculus*, *Rhododendron*, *Rosa*, *Rudbeckia*, *Senecio*, *Silene*, *Silphium*, *Sinapsis*, *Sor-*  
25 *bus*, *Spartium*, *Tecoma*, *Torenia*, *Tragopogon*, *Trollius*, *Tropaeolum*, *Tulipa*, *Tussilago*, *U-*  
*lex*, *Viola* oder *Zinnia* verwendet.
37. Verwendung der genetisch veränderten Organismen nach einem der Ansprüche 25 bis 36  
30 als Futter- oder Nahrungsmittel.
38. Verwendung der genetisch veränderten Organismen nach einem der Ansprüche 25 bis 36  
zur Herstellung von Ketocarotinoid-haltigen Extrakten oder zur Herstellung von Futter- und  
Nahrungsergänzungsmitteln.
- 25 39. Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 8 oder eine von dieser Se-  
quenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz,  
die eine Identität von mindestens 70 % auf Aminosäureebene mit der Sequenz SEQ. ID.  
NO. 8 aufweist, mit der Maßgabe, dass die Aminosäuresequenzen SEQ ID NO: 4 nicht ent-  
30 halten ist.
40. Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 6 oder eine von dieser Se-  
quenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz,  
die eine Identität von mindestens 70 % auf Aminosäureebene mit der Sequenz SEQ. ID.  
35 NO. 6 aufweist.
41. Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 12 oder eine von dieser Se-  
quenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz,  
die eine Identität von mindestens 70 % auf Aminosäureebene mit der Sequenz SEQ. ID.

NO. 12 aufweist, mit der Maßgabe, dass die Aminosäuresequenzen SEQ ID NO: 6 nicht enthalten ist.

- 5 42. Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 49 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 50 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 49 aufweist, mit der Maßgabe, dass die Aminosäuresequenzen SEQ ID NO: 47 nicht enthalten ist.
- 10 43. Nukleinsäure, kodierend ein Protein gemäß einem der Ansprüche 39 bis 42, mit der Maßgabe, dass die Sequenz SEQ ID NO: 5 nicht enthalten ist.
- 15 44. Verwendung eines Proteins, enthaltend die Aminosäuresequenz SEQ. ID. NO. 4 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 70 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 4 und die Eigenschaft einer Ketolase aufweist, als Ketolase.
- 20 45. Verwendung eines Proteins, enthaltend die Aminosäuresequenz SEQ. ID. NO. 6 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 65 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 6 und die Eigenschaft einer Ketolase aufweist, als Ketolase.
- 25 46. Verwendung eines Proteins, enthaltend die Aminosäuresequenz SEQ. ID. NO. 47 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 50 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 47 und die Eigenschaft einer Ketolase aufweist, als Ketolase.



## Verfahren zur Herstellung von Ketocarotinoiden in genetisch veränderten Organismen

## Beschreibung

- 5 Die vorliegende Erfindung betrifft ein Verfahren zur Herstellung von Ketocarotinoiden durch Kultivierung von genetisch veränderten Organismen, die im Vergleich zum Wildtyp eine veränderte Ketolase-Aktivität aufweisen, die genetisch veränderten Organismen, sowie deren Verwendung als Nahrungs- und Futtermittel und zur Herstellung von Ketocarotinoidextrakten.
- 10 Carotinoide werden de novo in Bakterien, Algen, Pilzen und Pflanzen synthetisiert. Ketocarotinoide, also Carotinoide, die mindestens eine Keto-Gruppe enthalten, wie beispielsweise Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin und Adonixanthin sind natürliche Antioxidantien und Pigmente, die von einigen Algen und Mikroorganismen als Sekundärmetabolite produziert werden.
- 15 Aufgrund ihrer farbgebenden Eigenschaften werden die Ketocarotinoide und insbesondere Astaxanthin als Pigmentierhilfsstoffe in der Tierernährung, insbesondere in der Forellen-, Lachs- und Shrimpszucht verwendet.
- 20 Die Herstellung von Astaxanthin erfolgt heutzutage größtenteils durch chemische Syntheseverfahren. Natürliche Ketocarotinoide, wie beispielsweise natürliches Astaxanthin, werden heutzutage in biotechnologischen Verfahren in kleinen Mengen durch Kultivierung von Algen, beispielsweise *Haematococcus pluvialis* oder durch Fermentation von gentechnologisch optimierten Mikroorganismen und anschließender Isolierung gewonnen.
- 25 Ein wirtschaftliches biotechnologisches Verfahren zur Herstellung von natürlichen Ketocarotinoiden ist daher von großer Bedeutung.
- 30 Nukleinsäuren kodierend eine Ketolase und die entsprechenden Proteinsequenzen sind aus verschiedenen Organismen isoliert und annotiert worden, wie beispielsweise Nukleinsäuren kodierend eine Ketolase aus *Agrobacterium aurantiacum* (EP 735 137, Accession NO: D58420), aus *Alcaligenes sp. PC-1* (EP 735137, Accession NO: D58422), *Haematococcus pluvialis* Flotow em. Wille und *Haematococcus pluvialis*, NIES-144 (EP 725137, WO 98/18910 und Lotan et al, FEBS Letters 1995, 364, 125-128, Accession NO: X86782 und D45881), *Paracoccus mar-*
- 35 *cusii* (Accession NO: Y15112), *Synechocystis sp. Strain PC6803* (Accession NO: NP\_442491), *Bradyrhizobium sp.* (Accession NO: AF218415) und *Nostoc sp. PCC 7120* (Kaneko et al, DNA Res. 2001, 8(5), 205 - 213; Accession NO: AP003592, BAB74888).

EP 735 137 beschreibt die Herstellung von Xanthophyllen in Mikroorganismen, wie beispielsweise *E. coli* durch Einbringen von Ketolase-Genen (*crtW*) aus *Agrobacterium aurantiacum* oder *Alcaligenes sp. PC-1* in Mikroorganismen.

- 5 Aus EP 725 137, WO 98/18910, Kajiwara et al. (Plant Mol. Biol. 1995, 29, 343-352) und Hirschberg et al. (FEBS Letters 1995, 364, 125-128) ist es bekannt, Astaxanthin durch Einbringen von Ketolase-Genen aus *Haematococcus pluvialis* (*crtW*, *crtO* oder *bkt*) in *E. coli* herzustellen.

- 10 Hirschberg et al. (FEBS Letters 1997, 404, 129-134) beschreiben die Herstellung von Astaxanthin in *Synechococcus* durch Einbringen von Ketolase-Genen (*crtO*) aus *Haematococcus pluvialis*. Sandmann et al. (Photochemistry and Photobiology 2001, 73(5), 551-55) beschreiben ein analoges Verfahren, das jedoch zur Herstellung von Canthaxanthin führt und nur Spuren Astaxanthin liefert.

- 15 WO 98/18910 und Hirschberg et al. (Nature Biotechnology 2000, 18(8), 888-892) beschreiben die Synthese von Ketocarotinoiden in Nektarien von Tabakblüten durch Einbringen des Ketolase-Gens aus *Haematococcus pluvialis* (*crtO*) in Tabak.

- 20 WO 01/20011 beschreibt ein DNA Konstrukt zur Produktion von Ketocarotinoiden, insbesondere Astaxanthin, in Samen von Ölsaatzpflanzen wie Raps, Sonnenblume, Sojabohne und Senf unter Verwendung eines Samen-spezifischen Promotors und einer Ketolase aus *Haematococcus pluvialis*.

- 25 Alle im Stand der Technik beschriebenen Verfahren zur Herstellung von Ketocarotinoiden und insbesondere die beschriebenen Verfahren zur Herstellung von Astaxanthin weisen den Nachteil auf, daß die transgenen Organismen eine große Menge an hydroxylierten Nebenprodukten, wie beispielsweise Zeaxanthin und Adonixanthin liefern.

- 30 Der Erfindung lag daher die Aufgabe zugrunde, ein Verfahren zur Herstellung von Ketocarotinoiden durch Kultivierung von genetisch veränderten Organismen zur Verfügung zu stellen, bzw. weitere genetisch veränderte Organismen, die Ketocarotinoide herstellen, zur Verfügung zu stellen, die die vorstehend beschriebenen Nachteile des Standes der Technik in geringerem Maße oder nicht mehr aufweisen.

- 35 Demgemäß wurde ein Verfahren zur Herstellung von Ketocarotinoiden gefunden, indem man genetisch veränderte Organismen kultiviert, die im Vergleich zum Wildtyp eine veränderte Ketolase-Aktivität aufweisen und die veränderte Ketolase-Aktivität durch eine Ketolase verursacht wird, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität

von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.

Die erfindungsgemäßen Organismen wie beispielsweise Mikroorganismen oder Pflanzen sind vorzugsweise als Ausgangsorganismen natürlicherweise in der Lage, Carotinoide wie beispielsweise  $\beta$ -Carotin oder Zeaxanthin herzustellen, oder können durch genetische Veränderung, wie beispielsweise Umregulierung von Stoffwechselwegen oder Komplementierung in die Lage versetzt werden, Carotinoide wie beispielsweise  $\beta$ -Carotin oder Zeaxanthin herzustellen.

Einige Organismen sind als Ausgangs- oder Wildtyporganismen bereits in der Lage, Ketocarotinoid wie beispielsweise Astaxanthin oder Canthaxanthin herzustellen. Diese Organismen, wie beispielsweise *Haematococcus pluvialis*, *Paracoccus marcusii*, *Xanthophyllomyces dendrorhous*, *Bacillus circulans*, *Chlorococcum*, *Phaffia rhodozyma*, *Adonisröschen*, *Neochloris wimmeri*, *Protosiphon botryoides*, *Scotiellopsis oocystiformis*, *Scenedesmus vacuolatus*, *Chlorella zofingiensis*, *Ankistrodesmus braunii*, *Euglena sanguinea*, *Bacillus atrophaeus*, *Blakeslea* weisen bereits als Ausgangs- oder Wildtyporganismus eine Ketolase-Aktivität auf.

In einer Ausführungsform des erfindungsgemäßen Verfahrens werden daher als Ausgangsorganismen Organismen verwendet, die bereits als Wildtyp oder Ausgangsorganismus eine Ketolaseaktivität aufweisen. In dieser Ausführungsform bewirkt die genetische Veränderung eine Erhöhung der Ketolase-Aktivität im Vergleich zum Wildtyp oder Ausgangsorganismus.

Unter Ketolase-Aktivität wird die Enzymaktivität einer Ketolase verstanden.

Unter einer Ketolase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am gegebenenfalls substituierten,  $\beta$ -Ionon-Ring von Carotinoiden eine Keto-Gruppe einzuführen.

Insbesondere wird unter einer Ketolase ein Protein verstanden, das die enzymatische Aktivität aufweist,  $\beta$ -Carotin in Canthaxanthin umzuwandeln.

Dementsprechend wird unter Ketolase-Aktivität die in einer bestimmten Zeit durch das Protein Ketolase umgesetzte Menge  $\beta$ -Carotin bzw. gebildete Menge Canthaxanthin verstanden.

Bei einer erhöhten Ketolase-Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein Ketolase die umgesetzte Menge  $\beta$ -Carotin bzw. die gebildete Menge Canthaxanthin erhöht.

Vorzugsweise beträgt diese Erhöhung der Ketolase-Aktivität mindestens 5 %, weiter bevorzugt mindestens 20 %, weiter bevorzugt mindestens 50 %, weiter bevorzugt mindestens 100 %, be-

vorzugter mindestens 300 %, noch bevorzugter mindestens 500 %, insbesondere mindestens 600 % der Ketolase-Aktivität des Wildtyps.

5 Unter dem Begriff "Wildtyp" wird erfindungsgemäß der entsprechende Ausgangsorganismus verstanden.

Je nach Zusammenhang kann unter dem Begriff "Organismus" der Ausgangsorganismus (Wildtyp) oder ein erfindungsgemäßer, genetisch veränderter Organismus oder beides verstanden werden.

10

Vorzugsweise und insbesondere in Fällen, in denen der Organismus oder der Wildtyp nicht eindeutig zugeordnet werden kann, wird unter "Wildtyp" für die Erhöhung oder Verursachung der Ketolase-Aktivität, für die nachstehend beschriebene Erhöhung der Hydroxylase-Aktivität, für die nachstehend beschriebene Erhöhung der  $\beta$ -Cyclase-Aktivität und die Erhöhung des Gehalts an Ketocarotinoiden jeweils ein Referenzorganismus verstanden.

15

Dieser Referenzorganismus ist für Mikroorganismen, die bereits als Wildtyp eine Ketolase Aktivität aufweisen, vorzugsweise *Haematococcus pluvialis*.

20 Dieser Referenzorganismus ist für Mikroorganismen, die als Wildtyp keine Ketolase Aktivität aufweisen, vorzugsweise *Blakeslea*.

Dieser Referenzorganismus ist für Pflanzen, die bereits als Wildtyp eine Ketolase-Aktivität aufweisen, vorzugsweise *Adonis aestivalis*, *Adonis flammeus* oder *Adonis annuus*, besonders bevorzugt *Adonis aestivalis*.

25

Dieser Referenzorganismus ist für Pflanzen, die als Wildtyp keine Ketolase-Aktivität in Blütenblätter aufweisen, vorzugsweise *Tagetes erecta*, *Tagetes patula*, *Tagetes lucida*, *Tagetes pringlei*, *Tagetes palmeri*, *Tagetes minuta* oder *Tagetes campanulata*, besonders bevorzugt *Tagetes erecta*.

30

Die Bestimmung der Ketolase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen und in Wildtyp- bzw. Referenzorganismen erfolgt vorzugsweise unter folgenden Bedingungen:

35

Die Bestimmung der Ketolase-Aktivität in Pflanzen- oder Mikroorganismenmaterial erfolgt in Anlehnung an die Methode von Frazer et al., (J. Biol. Chem. 272(10): 6128-6135, 1997). Die Ketolase-Aktivität in pflanzlichen oder Mikroorganismus-Extrakten wird mit den Substraten  $\beta$ -Carotin und Canthaxanthin in Gegenwart von Lipid (Sojalecithin) und Detergens (Natriumcholat)

bestimmt. Substrat/Produkt-Verhältnisse aus den Ketolase-Assays werden mittels HPLC ermittelt.

5 Die Erhöhung der Ketolase-Aktivität kann durch verschiedene Wege erfolgen, beispielsweise durch Ausschalten von hemmenden Regulationsmechanismen auf Translations- und Proteinebene oder durch Erhöhung der Genexpression einer Nukleinsäure, kodierend eine Ketolase, gegenüber dem Wildtyp, beispielsweise durch Induzierung des Ketolase-Gens durch Aktivatoren oder durch Einbringen von Nukleinsäuren, kodierend eine Ketolase, in den Organismus.

10 Unter Erhöhung der Genexpression einer Nukleinsäure, kodierend eine Ketolase, wird erfindungsgemäß in dieser Ausführungsform auch die Manipulation der Expression der Organismen eigenen endogenen Ketolasen verstanden. Dies kann beispielsweise durch Veränderung der Promotor DNA-Sequenz für Ketolase kodierende Gene erreicht werden. Eine solche Veränderung, die eine veränderte oder vorzugsweise erhöhte Expressionsrate mindestens eines endo-  
15 genen Ketolase Gens zur Folge hat, kann durch Deletion oder Insertion von DNA Sequenzen erfolgen.

Es ist wie vorstehend beschrieben möglich, die Expression mindestens einer endogenen Ketolase durch die Applikation exogener Stimuli zu verändern. Dies kann durch besondere physiologische Bedingungen, also durch die Applikation von Fremdstoffen erfolgen.  
20

Des weiteren kann eine erhöhte Expression mindestens eines endogenen Ketolase-Gens dadurch erzielt werden, dass ein im Wildtyporganismus nicht vorkommendes oder modifiziertes Regulatorprotein mit dem Promotor dieser Gene in Wechselwirkung tritt.  
25

Solch ein Regulator kann ein chimäres Protein darstellen, welches aus einer DNA-Bindedomäne und einer Transkriptionsaktivator-Domäne besteht, wie beispielsweise in WO 96/06166 beschrieben.  
30

In einer bevorzugten Ausführungsform erfolgt die Erhöhung der Ketolase-Aktivität gegenüber dem Wildtyp durch die Erhöhung der Genexpression einer Nukleinsäure, kodierend eine Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.  
35

In einer weiter bevorzugten Ausführungsform erfolgt die Erhöhung der Genexpression einer Nukleinsäure, kodierend eine Ketolase, durch Einbringen von Nukleinsäuren, die Ketolasen kodieren, in die Organismen, wobei die Ketolasen die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleite-

te Sequenz enthalten, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.

5 In den erfindungsgemäßen transgenen Organismen liegt also in dieser Ausführungsform gegenüber dem Wildtyp mindestens ein weiteres Ketolase-Gen vor, kodierend eine Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.

10 In dieser Ausführungsform weist der erfindungsgemäße genetisch veränderte Organismus dementsprechend mindestens eine exogene (=heterologe) Nukleinsäure, kodierend eine Ketolase, auf oder mindestens zwei endogene Nukleinsäuren, kodierend eine Ketolase, auf, wobei die Ketolasen die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz enthalten, die eine  
15 Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.

In einer anderen, bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens werden als Ausgangsorganismen Organismen verwendet, die als Wildtyp keine Ketolaseaktivität aufweisen.

20 In dieser bevorzugten Ausführungsform verursacht die genetische Veränderung die Ketolase-Aktivität in den Organismen. Der erfindungsgemäße genetisch veränderte Organismus weist somit in dieser bevorzugten Ausführungsform im Vergleich zum genetisch nicht veränderten Wildtyp eine Ketolase-Aktivität auf und ist somit vorzugsweise in der Lage, transgen eine Ketolase zu exprimieren, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser  
25 Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.

30 In dieser bevorzugten Ausführungsform erfolgt die Verursachung der Genexpression einer Nukleinsäure, kodierend eine Ketolase, analog zu der vorstehend beschriebenen Erhöhung der Genexpression einer Nukleinsäure, kodierend eine Ketolase, vorzugsweise durch Einbringen von Nukleinsäuren, die Ketolasen kodieren, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz  
35 SEQ. ID. NO. 2 aufweist, in den Ausgangsorganismus.

Dazu kann in beiden Ausführungsformen prinzipiell jede Nukleinsäuren, die eine Ketolase kodiert, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität

von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist, verwendet werden.

Die Verwendung der erfindungsgemäßen Nukleinsäuren, kodierend eine Ketolase, führt im erfindungsgemäßen Verfahren überraschenderweise zu Ketocarotinoiden mit einer geringeren Menge an hydroxylierten Nebenprodukten als bei der Verwendung der im Stand der Technik verwendeten Ketolase-Gene.

Alle in der Beschreibung erwähnten Nukleinsäuren können beispielsweise eine RNA-, DNA- oder cDNA-Sequenz sein.

Bei genomischen Ketolase-Sequenzen aus eukaryotischen Quellen, die Introns enthalten, sind für den Fall, dass der Wirtsorganismus nicht in der Lage ist oder nicht in die Lage versetzt werden kann, die entsprechenden Ketolase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen wie die entsprechenden cDNAs zu verwenden.

Beispiele für Nukleinsäuren, kodierend eine Ketolase, und die entsprechenden Ketolasen, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist, die im erfindungsgemäßen Verfahren vorteilhaft verwendet werden können, sind beispielsweise Sequenzen aus

*Nostoc sp. Strain PCC7120* (Accession NO: AP003592, BAB74888; Nukleinsäure: SEQ ID NO: 1, Protein SEQ ID NO: 2),

*Nostoc punctiforme ATTC 29133*, Nukleinsäure: Acc.-No. NZ\_AABC01000195, Basenpaar 55,604 bis 55,392 (SEQ ID NO: 3); Protein: Acc.-No. ZP\_00111258 (SEQ ID NO: 4) (als putatives Protein annotiert) oder

*Nostoc punctiforme ATTC 29133*, Nukleinsäure: Acc.-No. NZ\_AABC01000196, Basenpaar 140,571 bis 139,810 (SEQ ID NO: 5), Protein: (SEQ ID NO: 6) (nicht annotiert),

*Synechococcus sp. WH 8102*, Nukleinsäure: Acc.-No. NZ\_AABD01000001, Basenpaar 1,354,725-1,355,528 (SEQ ID NO: 46), Protein: Acc.-No. ZP\_00115639 (SEQ ID NO: 47) (als putatives Protein annotiert),

oder von diesen Sequenzen abgeleitete Ketolasesequenzen wie beispielsweise

die Ketolasen der Sequenz SEQ ID NO: 8 oder 10 und die entsprechenden kodierenden Nukleinsäuresequenzen SEQ ID NO: 7 oder SEQ ID NO: 9, die beispielsweise durch Variation/Mutation aus der Sequenz SEQ ID NO: 4 bzw. SEQ ID NO: 3 hervorgehen,

- 5 die Ketolasen der Sequenz SEQ ID NO: 12 oder 14 und die entsprechenden kodierenden Nukleinsäuresequenzen SEQ ID NO: 11 oder SEQ ID NO: 13, die beispielsweise durch Variation/Mutation aus der Sequenz SEQ ID NO: 6 bzw. SEQ ID NO: 5 hervorgehen, oder

- 10 die Ketolasen der Sequenz SEQ ID NO: 49 oder 51 und die entsprechenden kodierenden Nukleinsäuresequenzen SEQ ID NO: 48 oder SEQ ID NO: 50, die beispielsweise durch Variation bzw. Mutation aus der Sequenz SEQ ID NO: 47 bzw. SEQ ID NO: 46 hervorgehen.

- 15 Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene, die im erfindungsgemäßen Verfahren verwendet werden können, lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Aminosäuresequenzen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit der vorstehend beschriebenen Sequenzen SEQ ID NO: 2 leicht auffinden.

- 20 Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 1 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

- 25 Die Hybridisierung kann unter moderaten (geringe Stringenz) oder vorzugsweise unter stringenter (hohe Stringenz) Bedingungen erfolgen.

- 30 Solche Hybridisierungsbedingungen sind beispielsweise bei Sambrook, J., Fritsch, E.F., Maniatis, T., in: Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57 oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6 beschrieben.

- 35 Beispielhaft können die Bedingungen während des Waschschrilles ausgewählt sein aus dem Bereich von Bedingungen begrenzt von solchen mit geringer Stringenz (mit 2X SSC bei 50°C) und solchen mit hoher Stringenz (mit 0.2X SSC bei 50°C, bevorzugt bei 65°C) (20X SSC: 0,3 M Natriumcitrat, 3 M Natriumchlorid, pH 7.0).

Darüberhinaus kann die Temperatur während des Waschschrilles von moderaten Bedingungen bei Raumtemperatur, 22°C, bis zu stringenten Bedingungen bei 65°C angehoben werden.



Beide Parameter, Salzkonzentration und Temperatur, können gleichzeitig variiert werden, auch kann einer der beiden Parameter konstant gehalten und nur der andere variiert werden. Während der Hybridisierung können auch denaturierende Agenzien wie zum Beispiel Formamid oder SDS eingesetzt werden. In Gegenwart von 50 % Formamid wird die Hybridisierung bevorzugt bei 42°C ausgeführt.

Einige beispielhafte Bedingungen für Hybridisierung und Waschschrift sind infolge gegeben:

(1) Hybridisierungsbedingungen mit zum Beispiel

(i) 4X SSC bei 65°C, oder

(ii) 6X SSC bei 45°C, oder

(iii) 6X SSC bei 68°C, 100 mg/ml denaturierter Fischsperma-DNA, oder

(iv) 6X SSC, 0.5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachssperma-DNA bei 68°C, oder

(v) 6XSSC, 0.5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachssperma-DNA, 50 % Formamid bei 42°C, oder

(vi) 50 % Formamid, 4X SSC bei 42°C, oder

(vii) 50 % (vol/vol) Formamid, 0.1 % Rinderserumalbumin, 0.1 % Ficoll, 0.1 % Polyvinylpyrrolidon, 50 mM Natriumphosphatpuffer pH 6.5, 750 mM NaCl, 75 mM Natriumcitrat bei 42°C, oder

(viii) 2X oder 4X SSC bei 50°C (moderate Bedingungen), oder

(ix) 30 bis 40 % Formamid, 2X oder 4X SSC bei 42°C (moderate Bedingungen).

(2) Waschschrift für jeweils 10 Minuten mit zum Beispiel

(i) 0.015 M NaCl/0.0015 M Natriumcitrat/0.1 % SDS bei 50°C, oder

(ii) 0.1X SSC bei 65°C, oder

(iii) 0.1X SSC, 0.5 % SDS bei 68°C, oder

(iv) 0.1X SSC, 0.5 % SDS, 50 % Formamid bei 42°C, oder

5 (v) 0.2X SSC, 0.1 % SDS bei 42°C, oder

(vi) 2X SSC bei 65°C (moderate Bedingungen).

10 In einer bevorzugten Ausführungsform der erfindungsgemäßen Verfahren bringt man Nukleinsäuren ein, die eine Ketolase kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 50%, vorzugsweise mindestens 60%, vorzugsweise mindestens 65 %, vorzugsweise mindestens 70 %, bevorzugter mindestens 75 %, bevorzugter mindestens 80 %, bevorzugter mindestens 85 %, bevorzugter mindestens 90 %, bevorzugter mindestens 95 %, besonders bevorzugt mindestens 98 % auf Aminosäureebene mit der Sequenz SEQ ID NO: 2 aufweist.

20 Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die wie vorstehend beschrieben durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz, die ausgehend von der Sequenz SEQ ID NO: 2 durch künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

25 Unter dem Begriff "Substitution" ist in der Beschreibung der Austausch einer oder mehrerer Aminosäuren durch eine oder mehrere Aminosäuren zu verstehen. Bevorzugt werden sog. konservative Austausche durchgeführt, bei denen die ersetzte Aminosäure eine ähnliche Eigenschaft hat wie die ursprüngliche Aminosäure, beispielsweise Austausch von Glu durch Asp, Gln durch Asn, Val durch Ile, Leu durch Ile, Ser durch Thr.

30 Deletion ist das Ersetzen einer Aminosäure durch eine direkte Bindung. Bevorzugte Positionen für Deletionen sind die Termini des Polypeptides und die Verknüpfungen zwischen den einzelnen Proteindomänen.

35 Insertionen sind Einfügungen von Aminosäuren in die Polypeptidkette, wobei formal eine direkte Bindung durch ein oder mehrere Aminosäuren ersetzt wird.

Unter Identität zwischen zwei Proteinen wird die Identität der Aminosäuren über die jeweils gesamte Proteinlänge verstanden, insbesondere die Identität, die durch Vergleich mit Hilfe der Vector NTI Suite 7.1 Software der Firma Informax (USA) unter Anwendung der Clustal Methode

(Higgins DG, Sharp PM. Fast and sensitive multiple sequence alignments on a microcomputer. Comput Appl. Biosci. 1989 Apr;5(2):151-1) unter Einstellung folgender Parameter berechnet wird:

- 5 Multiple alignment parameter:
- |                                   |     |
|-----------------------------------|-----|
| Gap opening penalty               | 10  |
| Gap extension penalty             | 10  |
| Gap separation penalty range      | 8   |
| Gap separation penalty            | off |
| 10 % identity for alignment delay | 40  |
| Residue specific gaps             | off |
| Hydrophilic residue gap           | off |
| Transition weighing               | 0   |

- 15 Pairwise alignment parameter:

FAST algorithm on

K-tuple size 1

Gap penalty 3

Window size 5

- 20 Number of best diagonals 5

Unter einer Ketolase, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ ID NO: 2 aufweist, wird dementsprechend eine Ketolase verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 2, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 42 % aufweist.

25

Beispielsweise weist nach obigen Programmlogarithmus mit obigem Parametersatz die Sequenz der Ketolase aus *Nostoc punctiforme* ATTC 29133 (SEQ ID NO: 4) mit der Sequenz der Ketolase aus *Nostoc sp. Strain PCC7120* (SEQ ID NO: 2) eine Identität von 65% auf.

30

Die Sequenz der zweiten Ketolase aus *Nostoc punctiforme* ATTC 29133 (SEQ ID NO: 6) weist mit der Sequenz der Ketolase aus *Nostoc sp. Strain PCC7120* (SEQ ID NO: 2) beispielsweise eine Identität von 58% auf.

35

Die Sequenz der Ketolase aus *Synechococcus sp. WH 8102* (SEQ ID NO: 47) weist mit der Sequenz der Ketolase aus *Nostoc sp. Strain PCC7120* (SEQ ID NO: 2) beispielsweise eine Identität von 44% auf.

Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

- 5 Bevorzugt werden dafür solche Codons verwendet, die entsprechend der Organismus-spezifischen "codon usage" häufig verwendet werden. Die "codon usage" lässt sich anhand von Computerauswertungen anderer, bekannter Gene der betreffenden Organismen leicht ermitteln.

In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 1, in den Organismus ein.

10

Alle vorstehend erwähnten Ketolase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

15

20

Die Sequenz der Ketolase aus *Nostoc sp. Strain PCC7120* (SEQ ID NO: 2) weist mit den Sequenzen der Ketolasen die in den Verfahren des Standes der Technik verwendet werden eine Identität von 39% (*Agrobacterium aurantiacum* (EP 735 137, Accession NO: D58420), 40% (*Alcaligenes sp. PC-1* (EP 735137, Accession NO: D58422) und 20 bis 21 % (*Haematococcus pluvialis* Flotow em. Wille und *Haematococcus pluvialis*, NIES-144 (EP 725137, WO 98/18910 und Lotan et al, FEBS Letters 1995, 364, 125-128, Accession NO: X86782 und D45881) auf.

25

- 30 In einer bevorzugten Ausführungsform werden Organismen kultiviert, die gegenüber dem Wildtyp zusätzlich zur erhöhten Ketolase-Aktivität eine erhöhte Hydroxylase-Aktivität und/oder  $\beta$ -Cyclase-Aktivität aufweisen.

Unter Hydroxylase-Aktivität wird die Enzymaktivität einer Hydroxylase verstanden.

35

Unter einer Hydroxylase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten,  $\beta$ -Ionon-Ring von Carotinoiden eine Hydroxy-Gruppe einzuführen.

Insbesondere wird unter einer Hydroxylase ein Protein verstanden, das die enzymatische Aktivität aufweist,  $\beta$ -Carotin in Zeaxanthin oder Canthaxanthin in Astaxanthin umzuwandeln.

5 Dementsprechend wird unter Hydroxylase-Aktivität die in einer bestimmten Zeit durch das Protein Hydroxylase umgesetzte Menge  $\beta$ -Carotin oder Canthaxanthin bzw. gebildete Menge Zeaxanthin oder Astaxanthin verstanden.

10 Bei einer erhöhten Hydroxylase-Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein Hydroxylase die umgesetzte Menge  $\beta$ -Carotin oder Canthaxanthin bzw. die gebildete Menge Zeaxanthin oder Astaxanthin erhöht.

15 Vorzugsweise beträgt diese Erhöhung der Hydroxylase-Aktivität mindestens 5 %, weiter bevorzugt mindestens 20 %, weiter bevorzugt mindestens 50 %, weiter bevorzugt mindestens 100 %, bevorzugter mindestens 300 %, noch bevorzugter mindestens 500 %, insbesondere mindestens 600 % der Hydroxylase-Aktivität des Wildtyps.

Unter  $\beta$ -Cyclase-Aktivität wird die Enzymaktivität einer  $\beta$ -Cyclase verstanden.

20 Unter einer  $\beta$ -Cyclase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, einen endständigen, linearen Rest von Lycopin in einen  $\beta$ -Ionon-Ring zu überführen.

Insbesondere wird unter einer  $\beta$ -Cyclase ein Protein verstanden, das die enzymatische Aktivität aufweist,  $\gamma$ -Carotin in  $\beta$ -Carotin umzuwandeln.

25 Dementsprechend wird unter  $\beta$ -Cyclase-Aktivität die in einer bestimmten Zeit durch das Protein  $\beta$ -Cyclase umgesetzte Menge  $\gamma$ -Carotin bzw. gebildete Menge  $\beta$ -Carotin verstanden.

30 Bei einer erhöhten  $\beta$ -Cyclase -Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein  $\beta$ -Cyclase die umgesetzte Menge an Lycopin bzw.  $\gamma$ -Carotin oder die gebildete Menge an  $\gamma$ -Carotin aus Lycopin bzw. die gebildete Menge an  $\beta$ -Carotin aus  $\gamma$ -Carotin erhöht.

35 Vorzugsweise beträgt diese Erhöhung der  $\beta$ -Cyclase-Aktivität mindestens 5 %, weiter bevorzugt mindestens 20 %, weiter bevorzugt mindestens 50 %, weiter bevorzugt mindestens 100 %, bevorzugter mindestens 300 %, noch bevorzugter mindestens 500 %, insbesondere mindestens 600 % der  $\beta$ -Cyclase-Aktivität des Wildtyps.

Die Bestimmung der Hydroxylase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen und in Wildtyp- bzw. Referenzorganismen erfolgt vorzugsweise unter folgenden Bedingungen:

5 Die Aktivität der Hydroxylase wird nach Bouvier et al. (Biochim. Biophys. Acta 1391 (1998), 320-328) *in vitro* bestimmt. Es wird zu einer bestimmten Menge an Organismusextrakt Ferredoxin, Ferredoxin-NADP Oxidoreductase, Katalase, NADPH sowie  $\beta$ -Carotin mit Mono- und Digalaktosylglyzeriden zugegeben.

10 Besonders bevorzugt erfolgt die Bestimmung der Hydroxylase-Aktivität unter folgenden Bedingungen nach Bouvier, Keller, d'Harlingue und Camara (Xanthophyll biosynthesis: molecular and functional characterization of carotenoid hydroxylases from pepper fruits (*Capsicum annuum* L.; Biochim. Biophys. Acta 1391 (1998), 320-328):

15 Der *in-vitro* Assay wird in einem Volumen von 0.250 ml durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7.6), 0.025 mg Ferredoxin von Spinat, 0.5 Einheiten Ferredoxin-NADP+ Oxidoreduktase von Spinat, 0.25 mM NADPH, 0.010 mg  $\beta$ -Carotin (in 0.1 mg Tween 80 emulgiert), 0.05 mM einer Mischung von Mono- und Digalaktosylglyzeriden (1:1), 1 Einheit Katalase, 200 Mono- und Digalaktosylglyzeriden (1:1), 0.2 mg Rinderserumalbumin und Organismusextrakt in unterschiedlichem Volumen. Die Reaktionsmischung wird 2 Stunden bei 30°C inkubiert. Die Reaktionsprodukte werden mit organischem Lösungsmittel wie Aceton oder Chloroform/Methanol (2:1) extrahiert und mittels HPLC bestimmt.

25 Die Bestimmung der  $\beta$ -Cyclase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen und in Wildtyp- bzw. Referenzorganismen erfolgt vorzugsweise unter folgenden Bedingungen:

30 Die Aktivität der  $\beta$ -Cyclase wird nach Fraser und Sandmann (Biochem. Biophys. Res. Comm. 185(1) (1992) 9-15) *in vitro* bestimmt. Es werden zu einer bestimmten Menge an Organismusextrakt Kaliumphosphat als Puffer (pH 7.6), Lycopin als Substrat, Stromaprotein von Paprika, NADP+, NADPH und ATP zugegeben.

35 Besonders bevorzugt erfolgt die Bestimmung der  $\beta$ -Cyclase -Aktivität unter folgenden Bedingungen nach Bouvier, d'Harlingue und Camara (Molecular Analysis of carotenoid cyclase inhibition; Arch. Biochem. Biophys. 346(1) (1997) 53-64):

Der *in-vitro* Assay wird in einem Volumen von 250  $\mu$ l Volumen durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7.6), unterschiedliche Mengen an Organismusextrakt, 20 nM Lyco-

pin, 250 µg an chromoplastidärem Stromaprotein aus Paprika, 0.2 mM NADP<sup>+</sup>, 0.2 mM NADPH und 1 mM ATP. NADP/NADPH und ATP werden in 10 ml Ethanol mit 1 mg Tween 80 unmittelbar vor der Zugabe zum Inkubationsmedium gelöst. Nach einer Reaktionszeit von 60 Minuten bei 30°C wird die Reaktion durch Zugabe von Chloroform/Methanol (2:1) beendet. Die in Chloroform extrahierten Reaktionsprodukte werden mittels HPLC analysiert.

Ein alternativer Assay mit radioaktivem Substrat ist beschrieben in Fraser und Sandmann (Biochem. Biophys. Res. Comm. 185(1) (1992) 9-15).

10 Die Erhöhung der Hydroxylase-Aktivität und/oder  $\beta$ -Cyclase-Aktivität kann durch verschiedene Wege erfolgen, beispielsweise durch Ausschalten von hemmenden Regulationsmechanismen auf Expressions- und Proteinebene oder durch Erhöhung der Genexpression von Nukleinsäuren, kodierend eine Hydroxylase, und/oder von Nukleinsäuren, kodierend eine  $\beta$ -Cyclase, gegenüber dem Wildtyp.

15

Die Erhöhung der Genexpression der Nukleinsäuren, kodierend eine Hydroxylase, und/oder die Erhöhung der Genexpression der Nukleinsäure, kodierend eine  $\beta$ -Cyclase, gegenüber dem Wildtyp kann ebenfalls durch verschiedene Wege erfolgen, beispielsweise durch Induzierung des Hydroxylase-Gens und/oder  $\beta$ -Cyclase-Gens durch Aktivatoren oder durch Einbringen von einer oder mehrerer Hydroxylase-Genkopien und/oder  $\beta$ -Cyclase-Genkopien, also durch Einbringen mindestens einer Nukleinsäure, kodierend eine Hydroxylase, und/oder mindestens einer Nukleinsäure, kodierend eine  $\beta$ -Cyclase, in den Organismus.

20

Unter Erhöhung der Genexpression einer Nukleinsäure, kodierend eine Hydroxylase und/oder  $\beta$ -Cyclase, wird erfindungsgemäß auch die Manipulation der Expression der Organismus eigenen endogenen Hydroxylase und/oder  $\beta$ -Cyclase verstanden.

25

Dies kann beispielsweise durch Veränderung der Promotor DNA-Sequenz für Hydroxylasen und/oder  $\beta$ -Cyclasen kodierende Gene erreicht werden. Eine solche Veränderung, die eine erhöhte Expressionsrate des Gens zur Folge hat, kann beispielsweise durch Deletion oder Insertion von DNA Sequenzen erfolgen.

30

Es ist, wie vorstehend beschrieben, möglich, die Expression der endogenen Hydroxylase und/oder  $\beta$ -Cyclase durch die Applikation exogener Stimuli zu verändern. Dies kann durch besondere physiologische Bedingungen, also durch die Applikation von Fremdstoffen erfolgen.

35

Des weiteren kann eine veränderte bzw. erhöhte Expression eines endogenen Hydroxylase- und/oder  $\beta$ -Cyclase-Gens dadurch erzielt werden, dass ein im nicht transformierten Organismus nicht vorkommendes Regulator-Protein mit dem Promotor dieses Gens in Wechselwirkung tritt.

- 5 Solch ein Regulator kann ein chimäres Protein darstellen, welches aus einer DNA-Bindedomäne und einer Transkriptionsaktivator-Domäne besteht, wie beispielsweise in WO 96/06166 beschrieben.

- 10 In einer bevorzugten Ausführungsform erfolgt die Erhöhung der Genexpression einer Nukleinsäure, kodierend eine Hydroxylase, und/oder die Erhöhung der Genexpression einer Nukleinsäure, kodierend eine  $\beta$ -Cyclase, durch Einbringen von mindestens einer Nukleinsäure, kodierend eine Hydroxylase, und/oder durch Einbringen von mindestens einer Nukleinsäure, kodierend eine  $\beta$ -Cyclase, in den Organismus.

- 15 Dazu kann prinzipiell jedes Hydroxylase-Gen bzw. jedes  $\beta$ -Cyclase-Gen, also jede Nukleinsäure, die eine Hydroxylase und jede Nukleinsäure, die eine  $\beta$ -Cyclase kodiert, verwendet werden.

- 20 Bei genomischen Hydroxylase-bzw.  $\beta$ -Cyclase-Nukleinsäure-Sequenzen aus eukaryotischen Quellen, die Introns enthalten, sind für den Fall, dass der Wirtsorganismus nicht in der Lage ist oder nicht in die Lage versetzt werden kann, die entsprechende Hydroxylase bzw.  $\beta$ -Cyclase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs, zu verwenden.

- 25 Ein Beispiel für ein Hydroxylase-Gen ist eine Nukleinsäure, kodierend eine Hydroxylase, aus *Haematococcus pluvialis*, Accession AX038729, WO 0061764); (Nukleinsäure: SEQ ID NO: 15, Protein: SEQ ID NO: 16).

- 30 Ein Beispiel für ein  $\beta$ -Cyclase-Gen ist eine Nukleinsäure, kodierend eine  $\beta$ -Cyclase aus Tomate (Accession X86452). (Nukleinsäure: SEQ ID NO: 17, Protein: SEQ ID NO: 18).

- In den erfindungsgemäßen bevorzugten transgenen Organismen liegt also in dieser bevorzugten Ausführungsform gegenüber dem Wildtyp mindestens ein weiteres Hydroxylase-Gen und/oder  $\beta$ -Cyclase-Gen vor.

- 35 In dieser bevorzugten Ausführungsform weist der genetisch veränderte Organismus beispielsweise mindestens eine exogene Nukleinsäure, kodierend eine Hydroxylase, oder mindestens zwei endogene Nukleinsäuren, kodierend eine Hydroxylase und/oder mindestens eine exogene Nukleinsäure, kodierend eine  $\beta$ -Cyclase, oder mindestens zwei endogene Nukleinsäuren, kodie-



rend eine  $\beta$ -Cyclase, auf.

Bevorzugt verwendet man in vorstehend beschriebener bevorzugter Ausführungsform als Hydroxylase-Gene Nukleinsäuren, die Proteine kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 16 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 30 %, vorzugsweise mindestens 50 %, bevorzugter mindestens 70%, noch bevorzugter mindestens 90 %, am bevorzugtesten mindestens 95 % auf Aminosäureebene mit der Sequenz SEQ ID NO: 16, und die die enzymatische Eigenschaft einer Hydroxylase aufweisen.

Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, wie vorstehend beschrieben durch Homologievergleiche der Aminosäuresequenzen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit der SEQ ID. NO: 16 leicht auffinden.

Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich weiterhin beispielsweise ausgehend von der Sequenz SEQ ID NO: 15 aus verschiedenen Organismen deren genomische Sequenz nicht bekannt ist, wie vorstehend beschrieben, durch Hybridisierungs- und PCR-Techniken in an sich bekannter Weise leicht auffinden.

In einer weiter besonders bevorzugten Ausführungsform werden zur Erhöhung der Hydroxylase-Aktivität Nukleinsäuren in Organismen eingebracht, die Proteine kodieren, enthaltend die Aminosäuresequenz der Hydroxylase der Sequenz SEQ ID NO: 16.

Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

Bevorzugt werden dafür solche Kodons verwendet, die entsprechend des Organismus-spezifischen "codon usage" häufig verwendet werden. Dieser "codon usage" lässt sich anhand von Computerauswertungen anderer, bekannter Gene der betreffenden Organismen leicht ermitteln.

In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ. ID. NO: 15, in den Organismus ein.

Bevorzugt verwendet man in vorstehend beschriebener bevorzugter Ausführungsform als  $\beta$ -Cyclase-Gene Nukleinsäuren, die Proteine kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 18 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Amino-

säuren abgeleitete Sequenz, die eine Identität von mindestens 30 %, vorzugsweise mindestens 50 %, bevorzugter mindestens 70 %, noch bevorzugter mindestens 90 %, am bevorzugtesten mindestens 95 % auf Aminosäureebene mit der Sequenz SEQ ID NO: 18, und die die enzymatische Eigenschaft einer  $\beta$ -Cyclase aufweisen.

5

Weitere Beispiele für  $\beta$ -Cyclasen und  $\beta$ -Cyclase-Gene lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, wie vorstehend beschrieben durch Homologievergleiche der Aminosäuresequenzen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit der SEQ ID NO: 18 leicht auffinden.

10

Weitere Beispiele für  $\beta$ -Cyclasen und  $\beta$ -Cyclase-Gene lassen sich weiterhin beispielsweise ausgehend von der Sequenz SEQ ID NO: 17 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungs- und PCR-Techniken in an sich bekannter Weise leicht auffinden.

15

In einer weiter besonders bevorzugten Ausführungsform werden zur Erhöhung der  $\beta$ -Cyclase-Aktivität Nukleinsäuren in Organismen eingebracht, die Proteine kodieren, enthaltend die Aminosäuresequenz der  $\beta$ -Cyclase der Sequenz SEQ. ID. NO: 18.

20

Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

25

Bevorzugt werden dafür solche Kodons verwendet, die entsprechend des Organismus-spezifischen "codon usage" häufig verwendet werden. Dieser "codon usage" lässt sich anhand von Computerauswertungen anderer, bekannter Gene der betreffenden Organismen leicht ermitteln.

30

In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ. ID. NO: 17 in den Organismus ein.

35

Alle vorstehend erwähnten Hydroxylase-Gene oder  $\beta$ -Cyclase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoramiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A

laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

Besonders bevorzugt werden im erfindungsgemäßen Verfahren genetisch veränderte Organismen mit folgende Kombinationen genetischer Veränderungen verwendet:

5

Genetisch veränderte Organismen, die im Vergleich zum Wildtyp eine erhöhte oder verursachte Ketolase-Aktivität und eine erhöhte Hydroxylase-Aktivität aufweisen,

10

genetisch veränderte Organismen, die im Vergleich zum Wildtyp eine erhöhte oder verursachte Ketolase-Aktivität und eine erhöhte  $\beta$ -Cyclase-Aktivität aufweisen und

genetisch veränderte Organismen, die im Vergleich zum Wildtyp eine erhöhte oder verursachte Ketolase-Aktivität und eine erhöhte Hydroxylase-Aktivität und eine erhöhte  $\beta$ -Cyclase-Aktivität aufweisen.

15

Die Herstellung dieser genetisch veränderten Organismen kann, wie nachstehend beschrieben, beispielsweise durch Einbringen einzelner Nukleinsäurekonstrukte (Expressionskassetten) oder durch Einbringen von Mehrfachkonstrukten erfolgen, die bis zu zwei oder drei der beschriebenen Aktivitäten enthalten.

20

Unter Organismen werden erfindungsgemäß vorzugsweise Organismen verstanden, die als Wildtyp- oder Ausgangsorganismen natürlicherweise oder durch genetische Komplementierung und/oder Umregulierung der Stoffwechselwege in der Lage sind, Carotinoide, insbesondere  $\beta$ -Carotin und/oder Zeaxanthin und/oder Neoxanthin und/oder Violaxanthin und/oder Lutein herzustellen.

25

Weiter bevorzugte Organismen weisen als Wildtyp- oder Ausgangsorganismen bereits eine Hydroxylase-Aktivität auf und sind somit als Wildtyp- oder Ausgangsorganismen in der Lage, Zeaxanthin herzustellen.

30

Bevorzugte Organismen sind Pflanzen oder Mikroorganismen, wie beispielsweise Bakterien, Hefen, Algen oder Pilze.

35

Als Bakterien können sowohl Bakterien verwendet werden, die aufgrund des Einbringens von Genen der Carotinoidbiosynthese eines Carotinoid-produzierenden Organismus in der Lage sind, Xanthophylle zu synthetisieren, wie beispielsweise Bakterien der Gattung *Escherichia*, die beispielsweise crt-Gene aus *Erwinia* enthalten, als auch Bakterien, die von sich aus in der Lage sind, Xanthophylle zu synthetisieren wie beispielsweise Bakterien der Gattung *Erwinia*, *Agrobacterium*, *Flavobacterium*, *Alcaligenes*, *Paracoccus*, *Nostoc* oder Cyanobakterien der Gattung

*Synechocystis*.

- 5 Bevorzugte Bakterien sind *Escherichia coli*, *Erwinia herbicola*, *Erwinia uredovora*, *Agrobacterium aurantiacum*, *Alcaligenes* sp. PC-1, *Flavobacterium* sp. strain R1534, das Cyanobacterium *Synechocystis* sp. PCC6803, *Paracoccus marcusii* oder *Paracoccus carotinifaciens*.

Bevorzugte Hefen sind *Candida*, *Saccharomyces*, *Hansenula*, *Pichia* oder *Phaffia*. Besonders bevorzugte Hefen sind *Xanthophyllomyces dendrorhous* oder *Phaffia rhodozyma*.

- 10 Bevorzugte Pilze sind *Aspergillus*, *Trichoderma*, *Ashbya*, *Neurospora*, *Blakeslea*, *Phycomyces*, *Fusarium* oder weitere in Indian Chem. Engr. Section B. Vol. 37, No. 1, 2 (1995) auf Seite 15, Tabelle 6 beschriebene Pilze.

- 15 Bevorzugte Algen sind Grünalgen, wie beispielsweise Algen der Gattung *Haematococcus*, *Phaedactylum tricornatum*, *Volvox* oder *Dunaliella*. Besonders bevorzugte Algen sind *Haematococcus puvialis* oder *Dunaliella bardawil*.

- 20 Weitere brauchbare Mikroorganismen und deren Herstellung zur Durchführung des erfindungsgemäßen Verfahrens sind beispielsweise aus der DE-A-199 16 140 bekannt, worauf hiermit Bezug genommen wird.

- 25 Besonders bevorzugte Pflanzen sind Pflanzen ausgewählt aus den Familien Ranunculaceae, Berberidaceae, Papaveraceae, Cannabaceae, Rosaceae, Fabaceae, Linaceae, Vitaceae, Brassicaceae, Cucurbitaceae, Primulaceae, Caryophyllaceae, Amaranthaceae, Gentianaceae, Geraniaceae, Caprifoliaceae, Oleaceae, Tropaeolaceae, Solanaceae, Scrophulariaceae, Astera-

- 30 Ganz besonders bevorzugte Pflanzen sind ausgewählt aus der Gruppe der Pflanzengattungen *Marigold*, *Tagetes erecta*, *Tagetes patula*, *Acacia*, *Aconitum*, *Adonis*, *Arnica*, *Aquilegia*, *Aster*, *Astragalus*, *Bignonia*, *Calendula*, *Caltha*, *Campanula*, *Canna*, *Centaurea*, *Cheiranthus*, *Chrysanthemum*, *Citrus*, *Crepis*, *Crocus*, *Curcubita*, *Cytisus*, *Delonia*, *Delphinium*, *Dianthus*, *Dimorphotheca*, *Doronicum*, *Eschscholtzia*, *Forsythia*, *Fremontia*, *Gazania*, *Gelsemium*, *Genista*, *Gentiana*, *Geranium*, *Gerbera*, *Geum*, *Grevillea*, *Helenium*, *Helianthus*, *Hepatica*, *Heracleum*, *Hibiscus*, *Heliopsis*, *Hypericum*, *Hypochoeris*, *Impatiens*, *Iris*, *Jacaranda*, *Kerria*, *Laburnum*,  
35 *Lathyrus*, *Leontodon*, *Lilium*, *Linum*, *Lotus*, *Lycopersicon*, *Lysimachia*, *Marattia*, *Medicago*, *Mimulus*, *Narcissus*, *Oenothera*, *Osmanthus*, *Petunia*, *Photinia*, *Physalis*, *Phyteuma*, *Potentilla*, *Pyracantha*, *Ranunculus*, *Rhododendron*, *Rosa*, *Rudbeckia*, *Senecio*, *Silene*, *Silphium*, *Sinapsis*, *Sorbus*, *Spartium*, *Tecoma*, *Torenia*, *Tragopogon*, *Trollius*, *Tropaeolum*, *Tulipa*, *Tussilago*, *Ulex*, *Viola* oder *Zinnia*, besonders bevorzugt ausgewählt aus der Gruppe der Pflanzengattungen Ma-

rigold, Tagetes erecta, Tagetes patula, Lycopersicon, Rosa, Calendula, Physalis, Medicago, Helianthus, Chrysanthemum, Aster, Tulipa, Narcissus, Petunia, Geranium, Tropaeolum oder Adonis.

- 5 Im erfindungsgemäßen Verfahren zur Herstellung von Ketocarotinoiden wird vorzugsweise dem Kultivierungsschritt der genetisch veränderten Organismen ein Ernten der Organismen und weiter bevorzugt zusätzlich ein Isolieren von Ketocarotinoiden aus den Organismen angeschlossen.

- 10 Das Ernten der Organismen erfolgt in an sich bekannter Weise dem jeweiligen Organismus entsprechend. Mikroorganismen, wie Bakterien, Hefen, Algen oder Pilze oder Pflanzenzellen, die durch Fermentation in flüssigen Nährmedien kultiviert werden, können beispielsweise durch Zentrifugieren, Dekantieren oder Filtrieren abgetrennt werden. Pflanzen werden in an sich bekannter Weise auf Nährböden gezogen und entsprechend geerntet.

- 15 Die Kultivierung der genetisch veränderten Mikroorganismen erfolgt bevorzugt in Gegenwart von Sauerstoff bei einer Kultivierungstemperatur von mindestens etwa 20°C, wie z.B. 20°C bis 40°C, und einem pH-Wert von etwa 6 bis 9. Bei genetisch veränderten Mikroorganismen erfolgt vorzugsweise zunächst die Kultivierung der Mikroorganismen in Gegenwart von Sauerstoff und in einem Komplexmedium, wie z.B. TB- oder LB- Medium bei einer Kultivierungstemperatur von 20 etwa 20 °C oder mehr, und einem pH-Wert von etwa 6 bis 9, bis eine ausreichende Zelldichte erreicht ist. Um die Oxidationsreaktion besser steuern zu können, bevorzugt man die Verwendung eines induzierbaren Promotors. Die Kultivierung wird nach Induktion der Ketolaseexpression in Gegenwart von Sauerstoff, z.B. 12 Stunden bis 3 Tage, fortgesetzt.

- 25 Die Isolierung der Ketocarotinoide aus der geernteten Biomasse erfolgt in an sich bekannter Weise, beispielsweise durch Extraktion und gegebenenfalls weiterer chemische oder physikalischer Reinigungsprozesse, wie beispielsweise Fällungsmethoden, Kristallographie, thermische Trennverfahren, wie Rektifizierverfahren oder physikalische Trennverfahren, wie beispielsweise Chromatographie.

- 30 Wie nachstehend erwähnt, können die Ketocarotinoide in den erfindungsgemäßen, genetisch veränderten Pflanzen vorzugsweise in verschiedenen Pflanzengeweben, wie beispielsweise Samen, Blätter, Früchte, Blüten, insbesondere in Blütenblättern spezifisch hergestellt werden.

- 35 Die Isolierung von Ketocarotinoiden aus den geernteten Blütenblättern erfolgt in an sich bekannter Weise, beispielsweise durch Trocknung und anschließender Extraktion und gegebenenfalls weiterer chemischer oder physikalischer Reinigungsprozesse, wie beispielsweise Fällungsmethoden, Kristallographie, thermische Trennverfahren, wie Rektifizierverfahren oder physikalische Trennverfahren, wie beispielsweise Chromatographie. Die Isolierung von Ketocarotinoiden aus

den Blütenblättern erfolgt beispielsweise bevorzugt durch organische Lösungsmittel wie Aceton, Hexan, Ether oder tert.-Methylbutylether.

5 Weitere Isolierverfahren von Ketocarotinoiden, insbesondere aus Blütenblättern, sind beispielsweise in Egger und Kleinig (Phytochemistry (1967) 6, 437-440) und Egger (Phytochemistry (1965) 4, 609-618) beschrieben.

10 Vorzugsweise sind die Ketocarotinoide ausgewählt aus der Gruppe Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin und Adonixanthin.

Ein besonders bevorzugtes Ketocarotinoid ist Astaxanthin.

Je nach verwendetem Organismus fallen die Ketocarotinoide in freier Form oder als Fettsäureester an.

15 In Blütenblättern von Pflanzen fallen die Ketocarotinide im erfindungsgemäßen Verfahren in Form ihrer Mono- oder Diester mit Fettsäuren an. Einige nachgewiesene Fettsäuren sind z.B. Myristinsäure, Palmitinsäure, Stearinsäure, Ölsäure, Linolensäure, und Laurinsäure (Kamata und Simpson (1987) Comp. Biochem. Physiol. Vol. 86B(3), 587-591).

20 Die Herstellung der Ketocarotinoide kann in der ganzen Pflanze oder in einer bevorzugten Ausführungsform spezifisch in Pflanzengewebe, die Chromoplasten enthalten, erfolgen. Bevorzugte Pflanzengewebe sind beispielsweise Wurzeln, Samen, Blätter, Früchte, Blüten und insbesondere Nektarien und Blütenblätter, die auch Petalen bezeichnet werden.

25 In einer besonderen bevorzugten Ausführungsform der erfindungsgemäßen Verfahrens verwendet man genetisch veränderte Pflanzen, die in Blüten die höchste Expressionsrate einer Ketolase aufweisen.

30 Vorzugsweise wird dies dadurch erreicht, dass die Genexpression der Ketolase unter Kontrolle eines blütenspezifischen Promotors erfolgt. Beispielsweise werden dazu die vorstehend beschriebenen Nukleinsäuren, wie nachstehend ausführlich beschrieben, in einem Nukleinsäurekonstrukt funktionell verknüpft mit einem blütenspezifischen Promotor in die Pflanze eingebracht.

35 In einer weiteren, besonderen bevorzugten Ausführungsform der erfindungsgemäßen Verfahrens verwendet man genetisch veränderte Pflanzen, die in Früchten die höchste Expressionsrate einer Ketolase aufweisen.

Vorzugsweise wird dies dadurch erreicht, dass die Genexpression der Ketolase unter Kontrolle eines fruchtspezifischen Promotors erfolgt. Beispielsweise werden dazu die vorstehend beschriebenen Nukleinsäuren, wie nachstehend ausführlich beschrieben, in einem Nukleinsäurekonstrukt funktionell verknüpft mit einem fruchtspezifischen Promotor in die Pflanze eingebracht.

5

In einer weiteren, besonders bevorzugten, Ausführungsform der erfindungsgemäßen Verfahrens verwendet man genetisch veränderte Pflanzen, die in Samen die höchste Expressionsrate einer Ketolase aufweisen.

10

Vorzugsweise wird dies dadurch erreicht, dass die Genexpression der Ketolase unter Kontrolle eines samenspezifischen Promotors erfolgt. Beispielsweise werden dazu die vorstehend beschriebenen Nukleinsäuren, wie nachstehend ausführlich beschrieben, in einem Nukleinsäurekonstrukt funktionell verknüpft mit einem samenspezifischen Promotor in die Pflanze eingebracht.

5

Das Targeting in die Chromoplasten erfolgt durch ein funktionell verknüpftes plastidäres Transitpeptid.

20

Im folgenden wird exemplarisch die Herstellung genetisch veränderter Pflanzen mit erhöhter oder verursachter Ketolase-Aktivität beschrieben. Die Erhöhung weiterer Aktivitäten, wie beispielsweise der Hydroxylase-Aktivität und/oder der  $\beta$ -Cyclase-Aktivität kann analog unter Verwendung von Nukleinsäuresequenzen, kodierend eine Hydroxylase bzw.  $\beta$ -Cyclase anstelle von Nukleinsäuresequenzen, kodierend eine Ketolase, erfolgen. Die Transformation kann bei den Kombinationen von genetischen Veränderungen einzeln oder durch Mehrfachkonstrukte erfolgen.

25

30

Die Herstellung der transgenen Pflanzen erfolgt vorzugsweise durch Transformation der Ausgangspflanzen, mit einem Nukleinsäurekonstrukt, das die vorstehend beschriebenen Nukleinsäuren, kodierend eine Ketolase enthält, die mit einem oder mehreren Regulationssignalen funktionell verknüpft sind, die die Transkription und Translation in Pflanzen gewährleisten.

35

Diese Nukleinsäurekonstrukte, in denen die kodierende Nukleinsäuresequenz mit einem oder mehreren Regulationssignalen funktionell verknüpft sind, die die Transkription und Translation in Pflanzen gewährleisten, werden im folgenden auch Expressionskassetten genannt.

Vorzugsweise enthalten die Regulationssignale einen oder mehrere Promotoren, die die Transkription und Translation in Pflanzen gewährleisten.

Die Expressionskassetten beinhalten Regulationssignale, also regulative Nukleinsäuresequenzen, welche die Expression der kodierenden Sequenz in der Wirtszelle steuern. Gemäß einer bevorzugten Ausführungsform umfasst eine Expressionskassette stromaufwärts, d.h. am 5'-Ende der kodierenden Sequenz, einen Promotor und stromabwärts, d.h. am 3'-Ende, ein Polyadenylierungssignal und gegebenenfalls weitere regulatorische Elemente, welche mit der dazwischenliegenden kodierenden Sequenz für mindestens eines der vorstehend beschriebenen Gene operativ verknüpft sind. Unter einer operativen Verknüpfung versteht man die sequenzielle Anordnung von Promotor, kodierender Sequenz, Terminator und ggf. weiterer regulativer Elemente derart, das jedes der regulativen Elemente seine Funktion bei der Expression der kodierenden Sequenz bestimmungsgemäß erfüllen kann.

Im folgenden werden beispielhaft die bevorzugten Nukleinsäurekonstrukte, Expressionskassetten und Vektoren für Pflanzen und Verfahren zur Herstellung von transgenen Pflanzen, sowie die transgenen Pflanzen selbst beschrieben.

Die zur operativen Verknüpfung bevorzugten, aber nicht darauf beschränkten Sequenzen, sind Targeting-Sequenzen zur Gewährleistung der subzellulären Lokalisation im Apoplasten, in der Vakuole, in Plastiden, im Mitochondrium, im Endoplasmatischen Retikulum (ER), im Zellkern, in Ölkörperchen oder anderen Kompartimenten und Translationsverstärkern wie die 5'-Führungssequenz aus dem Tabak-Mosaik-Virus (Gallie et al., Nucl. Acids Res. 15 (1987), 8693 -8711).

Als Promotor der Expressionskassette ist grundsätzlich jeder Promotor geeignet, der die Expression von Fremdgenen in Pflanzen steuern kann.

"Konstitutiver" Promotor meint solche Promotoren, die eine Expression in zahlreichen, bevorzugt allen, Geweben über einen größeren Zeitraum der Pflanzenentwicklung, bevorzugt zu allen Zeitpunkten der Pflanzenentwicklung, gewährleisten.

Vorzugsweise verwendet man insbesondere einen pflanzlichen Promotor oder einen Promotor, der einem Pflanzenvirus entstammt. Insbesondere bevorzugt ist der Promotor des 35S-Transkriptes des CaMV Blumenkohlmosaikvirus (Franck et al. (1980) Cell 21:285-294; Odell et al. (1985) Nature 313:810-812; Shewmaker et al. (1985) Virology 140:281-288; Gardner et al. (1986) Plant Mol Biol 6:221-228), der 19S CaMV Promotor (US 5,352,605; WO 84/02913; Benfey et al. (1989) EMBO J 8:2195-2202), den Triose-Phosphat Translokator (TPT) Promotor aus *Arabidopsis thaliana* Acc.-No. AB006698, Basenpaar 53242 bis 55281; das Gen beginnend ab bp 55282 ist mit "phosphate/triose-phosphate translocator" annotiert, oder den 34S Promoter aus Figwort mosaic virus Acc.-No. X16673, Basenpaar 1 bis 554.



Ein weiterer geeigneter konstitutiver Promotor ist der pds Promoter (Pecker et al. (1992) Proc. Natl. Acad. Sci USA 89: 4962-4966) oder der "Rubisco small subunit (SSU)"-Promotor (US 4,962,028), der LeguminB-Promotor (GenBank Acc.-Nr. X03677), der Promotor der Nopalinsynthase aus Agrobacterium, der TR-Doppelpromotor, der OCS (Octopin Synthase) Promotor aus Agrobacterium, der Ubiquitin Promotor (Holtorf S et al. (1995) Plant Mol Biol 29:637-649),  
5 der Ubiquitin 1 Promotor (Christensen et al. (1992) Plant Mol Biol 18:675-689; Bruce et al. (1989) Proc Natl Acad Sci USA 86:9692-9696), der Smas Promotor, der Cinnamylalkoholdehydrogenase-Promotor (US 5,683,439), die Promotoren der vakuolärer ATPase Untereinheiten oder der Promotor eines prolinreichen Proteins aus Weizen (WO 91/13991), der Pnit-Promoter  
10 (Y07648.L, Hillebrand et al. (1998), Plant. Mol. Biol. 36, 89-99, Hillebrand et al. (1996), Gene, 170, 197-200) sowie weitere Promotoren von Genen, deren konstitutive Expression in Pflanzen dem Fachmann bekannt ist.

Die Expressionskassetten können auch einen chemisch induzierbaren Promotor enthalten (Übersichtsartikel: Gatz et al. (1997) Annu Rev Plant Physiol Plant Mol Biol 48:89-108), durch den  
5 die Expression des Ketolase-Gens in der Pflanze zu einem bestimmten Zeitpunkt gesteuert werden kann. Derartige Promotoren, wie z.B. der PRP1 Promotor (Ward et al. (1993) Plant Mol Biol 22:361-366), ein durch Salicylsäure induzierbarer Promotor (WO 95/19443), ein durch Benzolsulfonamid-induzierbarer Promotor (EP 0 388 186), ein durch Tetrazyklin-induzierbarer Promotor (Gatz et al. (1992) Plant J 2:397-404), ein durch Abscisinsäure induzierbarer Promotor  
20 (EP 0 335 528) bzw. ein durch Ethanol- oder Cyclohexanon-induzierbarer Promotor (WO 93/21334) können ebenfalls verwendet werden.

Ferner sind Promotoren bevorzugt, die durch biotischen oder abiotischen Stress induziert werden wie beispielsweise der pathogen-induzierbare Promotor des PRP1-Gens (Ward et al. (1993) Plant Mol Biol 22:361-366), der hitzeinduzierbare hsp70- oder hsp80-Promoter aus Tomate (US  
25 5,187,267), der kälteinduzierbare alpha-Amylase Promoter aus der Kartoffel (WO 96/12814), der licht-induzierbare PPK Promotor oder der verwundungsinduzierte pinII-Promoter (EP375091).

Pathogen-induzierbare Promotoren umfassen die von Genen, die infolge eines Pathogenbefalls induziert werden wie beispielsweise Gene von PR-Proteinen, SAR-Proteinen, b-1,3-Glucanase, Chitinase usw. (beispielsweise Redolfi et al. (1983) Neth J Plant Pathol 89:245-254; Uknes, et al. (1992) The Plant Cell 4:645-656; Van Loon (1985) Plant Mol Biol 4:111-116; Marineau et al.  
30 (1987) Plant Mol Biol 9:335-342; Matton et al. (1987) Molecular Plant-Microbe Interactions 2:325-342; Somssich et al. (1986) Proc Natl Acad Sci USA 83:2427-2430; Somssich et al. (1988) Mol Gen Genetics 2:93-98; Chen et al. (1996) Plant J 10:955-966; Zhang and Sing (1994) Proc Natl Acad Sci USA 91:2507-2511; Warner, et al. (1993) Plant J 3:191-201; Siebertz

et al. (1989) Plant Cell 1:961-968(1989).

Umfasst sind auch verwundungsinduzierbare Promotoren wie der des pinII-Gens (Ryan (1990) Ann Rev Phytopath 28:425-449; Duan et al. (1996) Nat Biotech 14:494-498), des wun1 und  
5 wun2-Gens (US 5,428,148), des win1- und win2-Gens (Stanford et al. (1989) Mol Gen Genet 215:200-208), des Systemin-Gens (McGurl et al. (1992) Science 225:1570-1573), des WIP1-Gens (Rohmeier et al. (1993) Plant Mol Biol 22:783-792; Ekelkamp et al. (1993) FEBS Letters 323:73-76), des MPI-Gens (Corderok et al. (1994) The Plant J 6(2):141-150) und dergleichen.

10 Weitere geeignete Promotoren sind beispielsweise fruchtreifung-spezifische Promotoren, wie beispielsweise der fruchtreifung-spezifische Promotor aus Tomate (WO 94/21794, EP 409 625). Entwicklungsabhängige Promotoren schließt zum Teil die gewebespezifischen Promotoren ein, da die Ausbildung einzelner Gewebe naturgemäß entwicklungsabhängig erfolgt.

15 Weiterhin sind insbesondere solche Promotoren bevorzugt, die die Expression in Geweben oder Pflanzenteilen sicherstellen, in denen beispielsweise die Biosynthese von Ketocarotinoiden bzw. dessen Vorstufen stattfindet. Bevorzugt sind beispielsweise Promotoren mit Spezifitäten für die Antheren, Ovarien, Petalen, Sepalen, Blüten, Blätter, Stengel, Samen und Wurzeln und Kombinationen hieraus.

20 Knollen-, Speicherwurzel- oder Wurzel-spezifische Promotoren sind beispielsweise der Patatin-Promotor Klasse I (B33) oder der Promotor des Cathepsin D Inhibitors aus Kartoffel.

25 Blattspezifische Promotoren sind beispielsweise der Promotor der cytosolischen FBPase aus Kartoffel (WO 97/05900), der SSU Promotor (small subunit) der Rubisco (Ribulose-1,5-bisphosphatcarboxylase) oder der ST-LSI Promotor aus Kartoffel (Stockhaus et al. (1989) EMBO J 8:2445-2451).

30 Blüten-spezifische Promotoren sind beispielsweise der Phytoen-Synthase Promotor (WO 92/16635) oder der Promotor des P-rr Gens (WO 98/22593), der AP3 Promoter aus Arabidopsis thaliana (siehe Beispiel 5), der CHRC-Promoter (Chromoplast-specific carotenoid-associated protein (CHRC) gene promoter aus Cucumis sativus Acc.-No. AF099501, Basenpaar 1 bis 1532), der EPSP\_Synthase Promotor (5-enolpyruvylshikimate-3-phosphate synthase gene promoter aus Petunia hybrida, Acc.-No. M37029, Basenpaar 1 bis 1788), der PDS Promotor  
35 (Phytoene desaturase gene promoter aus Solanum lycopersicum, Acc.-No. U46919, Basenpaar 1 bis 2078), der DFR-A Promotor (Dihydroflavonol 4-reductase gene A promoter aus Petunia hybrida, Acc.-No. X79723, Basenpaar 32 bis 1902) oder der FBP1 Promotor (Floral Binding Protein 1 gene promoter aus Petunia hybrida, Acc.-No. L10115, Basenpaar 52 bis 1069).

Antheren-spezifische Promotoren sind beispielsweise der 5126-Promotor (US 5,689,049, US 5,689,051), der glob-I Promotor oder der g-Zein Promotor.

5 Samen-spezifische Promotoren sind beispielsweise der ACP05-Promotor (Acyl-carrier-Protein Gen, WO9218634), die Promotoren AtS1 und AtS3 von *Arabidopsis* (WO 9920775), der LeB4-Promotor von *Vicia faba* (WO 9729200 und US 06403371), der Napin-Promotor von *Brassica napus* (US 5608152; EP 255378; US 5420034), der SBP-Promotor von *Vicia faba* (DE 9903432) oder die Maispromotoren End1 und End2 (WO 0011177).

10 Weitere zur Expression in Pflanzen geeignete Promotoren sind beschrieben in Rogers et al. (1987) Meth in Enzymol 153:253-277; Schardl et al. (1987) Gene 61:1-11 und Berger et al. (1989) Proc Natl Acad Sci USA 86:8402-8406).

15 Besonders bevorzugt im erfindungsgemäßen Verfahren sind konstitutive, samenspezifische, fruchtspezifische, blütenspezifische und insbesondere blütenblattspezifische Promotoren.

20 Die vorliegende Erfindung betrifft daher insbesondere ein Nukleinsäurekonstrukt, enthaltend funktionell verknüpft einen blütenspezifischen oder insbesondere einen blütenblattspezifischen Promotor und eine Nukleinsäure, kodierend eine Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.

25 Die Herstellung einer Expressionskassette erfolgt vorzugsweise durch Fusion eines geeigneten Promotors mit einer vorstehend beschriebenen Nukleinsäure, kodierend eine Ketolase, und vorzugsweise einer zwischen Promotor und Nukleinsäure-Sequenz inserierten Nukleinsäure, die für ein plastidenspezifisches Transitpeptid kodiert, sowie einem Polyadenylierungssignal nach gängigen Rekombinations- und Klonierungstechniken, wie sie beispielsweise in T. Maniatis, E.F. Fritsch und J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in T.J. Silhavy, M.L. Berman und L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience (1987), beschrieben sind.

35 Die vorzugsweise inserierte Nukleinsäuren, kodierend ein plastidäres Transitpeptid, gewährleisten die Lokalisation in Plastiden und insbesondere in Chromoplasten.

Es können auch Expressionskassetten verwendet werden, deren Nukleinsäure-Sequenz für ein Ketolase-Fusionsprotein kodiert, wobei ein Teil des Fusionsproteins ein Transitpeptid ist, das

die Translokation des Polypeptides steuert. Bevorzugt sind für die Chromoplasten spezifische Transitpeptide, welche nach Translokation der Ketolase in die Chromoplasten vom Ketolase-Teil enzymatisch abgespalten werden.

- 5 Insbesondere bevorzugt ist das Transitpeptid, das von der plastidären *Nicotiana tabacum* Transketolase oder einem anderen Transitpeptid (z.B. dem Transitpeptid der kleinen Untereinheit der Rubisco (rbcS) oder der Ferredoxin NADP Oxidoreduktase als auch der Isopentenylpyrophosphat Isomerase-2) oder dessen funktionellem Äquivalent abgeleitet ist.
- 10 Besonders bevorzugt sind Nukleinsäure-Sequenzen von drei Kassetten des Plastiden-Transitpeptids der plastidären Transketolase aus Tabak in drei Leserastern als KpnI/BamHI Fragmente mit einem ATG-Codon in der NcoI Schnittstelle:

pTP09

15

KpnI\_GGTACCATGGCGTCTTCTTCTTCTCACTCTCTCTCAAGCTATCCTCTCTCGTTCTG  
TCCCTCGCCATGGCTCTGCCTCTTCTTCTCAACTTTCCCCTTCTTCTCT-  
CACTTTTTCCGGCCTTAAATCCAATCCCAATATCACCACCTCCCGCCGCCG-  
TACTCCTTCCTCCGCCGCCGCCGCCGCCGTCGTAAGGTCACCGGCGATTCTGCCT-  
20 CAGCTGCAACCGAAACCATAGAGAAACTGAGACTGCGGGATCC\_BamHI

pTP10

25

KpnI\_GGTACCATGGCGTCTTCTTCTTCTCACTCTCTCTCAAGCTATCCTCTCTCGTTCTG  
TCCCTCGCCATGGCTCTGCCTCTTCTTCTCAACTTTCCCCTTCTTCTCT-  
CACTTTTTCCGGCCTTAAATCCAATCCCAATATCACCACCTCCCGCCGCCG-  
TACTCCTTCCTCCGCCGCCGCCGCCGCCGTCGTAAGGTCACCGGCGATTCTGCCT-  
CAGCTGCAACCGAAACCATAGAGAAACTGAGACTGCGCTGGATCC\_BamHI

30 pTP11

35

KpnI\_GGTACCATGGCGTCTTCTTCTTCTCACTCTCTCTCAAGCTATCCTCTCTCGTTCTG  
TCCCTCGCCATGGCTCTGCCTCTTCTTCTCAACTTTCCCCTTCTTCTCT-  
CACTTTTTCCGGCCTTAAATCCAATCCCAATATCACCACCTCCCGCCGCCG-  
TACTCCTTCCTCCGCCGCCGCCGCCGCCGTCGTAAGGTCACCGGCGATTCTGCCT-  
CAGCTGCAACCGAAACCATAGAGAAACTGAGACTGCGGGATCC\_BamHI

Weitere Beispiele für ein plastidäres Transitpeptid sind das Transitpeptid der plastidären Isopentenyl-pyrophosphat Isomerase-2 (IPP-2) aus *Arabidopsis thaliana* und das Transitpeptid der

kleinen Untereinheit der Ribulosebisphosphat Carboxylase (rbcS) aus Erbse (Guerineau, F, Woolston, S, Brooks, L, Mullineaux, P (1988) An expression cassette for targeting foreign proteins into the chloroplasts. Nucl. Acids Res. 16: 11380).

- 5 Die erfindungsgemäßen Nukleinsäuren können synthetisch hergestellt oder natürlich gewonnen sein oder eine Mischung aus synthetischen und natürlichen Nukleinsäure-Bestandteilen enthalten, sowie aus verschiedenen heterologen Genabschnitten verschiedener Organismen bestehen.
- 10 Bevorzugt sind, wie vorstehend beschrieben, synthetische Nukleotid-Sequenzen mit Kodons, die von Pflanzen bevorzugt werden. Diese von Pflanzen bevorzugten Kodons können aus Kodons mit der höchsten Proteinhäufigkeit bestimmt werden, die in den meisten interessanten Pflanzenspezies exprimiert werden.
- 15 Bei der Präparation einer Expressionskassette können verschiedene DNA-Fragmente manipuliert werden, um eine Nukleotid-Sequenz zu erhalten, die zweckmäßigerweise in der korrekten Richtung liest und die mit einem korrekten Leseraster ausgestattet ist. Für die Verbindung der DNA-Fragmente miteinander können an die Fragmente Adaptoren oder Linker angesetzt werden.
- 20 Zweckmäßigerweise können die Promotor- und die Terminator-Regionen in Transkriptionsrichtung mit einem Linker oder Polylinker, der eine oder mehrere Restriktionsstellen für die Insertion dieser Sequenz enthält, versehen werden. In der Regel hat der Linker 1 bis 10, meistens 1 bis 8, vorzugsweise 2 bis 6 Restriktionsstellen. Im allgemeinen hat der Linker innerhalb der regulatorischen Bereiche eine Größe von weniger als 100 bp, häufig weniger als 60 bp, mindestens jedoch 5 bp. Der Promotor kann sowohl nativ bzw. homolog als auch fremdartig bzw. heterolog zur Wirtspflanze sein. Die Expressionskassette beinhaltet vorzugsweise in der 5'-3'-
- 25 Transkriptionsrichtung den Promotor, eine kodierende Nukleinsäuresequenz oder ein Nukleinsäurekonstrukt und eine Region für die transkriptionale Termination. Verschiedene Terminationsbereiche sind gegeneinander beliebig austauschbar.
- 30 Beispiele für einen Terminator sind der 35S-Terminator (Guerineau et al. (1988) Nucl Acids Res. 16: 11380), der nos Terminator (Depicker A, Stachel S, Dhaese P, Zambryski P, Goodman HM. Nopaline synthase: transcript mapping and DNA sequence. J Mol Appl Genet. 1982;1(6):561-73) oder der ocs Terminator (Gielen, J, de Beuckeleer, M, Seurinck, J, Debroek, H, de Greve, H, Lemmers, M, van Montagu, M, Schell, J (1984) The complete sequence of the TL-DNA of the Agrobacterium tumefaciens plasmid pTiAch5. EMBO J. 3: 835-846).
- 35

Ferner können Manipulationen, die passende Restriktionsschnittstellen bereitstellen oder die überflüssige DNA oder Restriktionsschnittstellen entfernen, eingesetzt werden. Wo Insertionen, Deletionen oder Substitutionen wie z.B. Transitionen und Transversionen in Frage kommen, können *in vitro*-Mutagenese, "primer-repair", Restriktion oder Ligation verwendet werden.

5

Bei geeigneten Manipulationen, wie z.B. Restriktion, "chewing-back" oder Auffüllen von Überhängen für "bluntends", können komplementäre Enden der Fragmente für die Ligation zur Verfügung gestellt werden.

10

Bevorzugte Polyadenylierungssignale sind pflanzliche Polyadenylierungssignale, vorzugsweise solche, die im wesentlichen T-DNA-Polyadenylierungssignale aus *Agrobacterium tumefaciens*, insbesondere des Gens 3 der T-DNA (Octopin Synthase) des Ti-Plasmids pTiACH5 entsprechen (Gielen et al., EMBO J. 3 (1984), 835 ff) oder funktionelle Äquivalente.

15

Die Übertragung von Fremdgenen in das Genom einer Pflanze wird als Transformation bezeichnet.

Dazu können an sich bekannte Methoden zur Transformation und Regeneration von Pflanzen aus Pflanzengeweben oder Pflanzenzellen zur transienten oder stabilen Transformation genutzt werden.

20

25

Geeignete Methoden zur Transformation von Pflanzen sind die Protoplastentransformation durch Polyethylenglykol-induzierte DNA-Aufnahme, das biolistische Verfahren mit der Genkanone – die sogenannte "particle bombardment" Methode, die Elektroporation, die Inkubation trockener Embryonen in DNA-haltiger Lösung, die Mikroinjektion und der, vorstehend beschriebene, durch *Agrobacterium* vermittelte Gentransfer. Die genannten Verfahren sind beispielsweise in B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu, Academic Press (1993), 128-143 sowie in Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225) beschrieben.

30

Vorzugsweise wird das zu exprimierende Konstrukt in einen Vektor kloniert, der geeignet ist, *Agrobacterium tumefaciens* zu transformieren, beispielsweise pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984), 8711) oder besonders bevorzugt pSUN2, pSUN3, pSUN4 oder pSUN5 (WO 02/00900).

35

Mit einem Expressionsplasmid transformierte Agrobakterien können in bekannter Weise zur Transformation von Pflanzen verwendet werden, z.B. indem verwundete Blätter oder Blattstücke in einer Agrobakterienlösung gebadet und anschließend in geeigneten Medien kultiviert werden.

Zur bevorzugten Herstellung von genetisch veränderten Pflanzen, im folgenden auch transgene Pflanzen bezeichnet, wird die fusionierte Expressionskassette, die eine Ketolase exprimiert, in einen Vektor, beispielsweise pBin19 oder insbesondere pSUN5 und pSUN3 kloniert, der geeignet ist, in *Agrobacterium tumefaciens* transformiert zu werden. Mit einem solchen Vektor transformierte Agrobakterien können dann in bekannter Weise zur Transformation von Pflanzen, insbesondere von Kulturpflanzen verwendet werden, indem beispielsweise verwundete Blätter oder Blattstücke in einer Agrobakterienlösung gebadet und anschließend in geeigneten Medien kultiviert werden.

- 10 Die Transformation von Pflanzen durch Agrobakterien ist unter anderem bekannt aus F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu, Academic Press, 1993, S. 15-38. Aus den transformierten Zellen der verwundeten Blätter bzw. Blattstücke können in bekannter Weise transgene Pflanzen regeneriert werden, die ein in die Expressionskassette integriertes Gen für die Expression einer Nukleinsäure, kodierend eine Ketolase, enthalten.

- 15 Zur Transformation einer Wirtspflanze mit einer für eine Ketolase kodierenden Nukleinsäure wird eine Expressionskassette als Insertion in einen rekombinanten Vektor eingebaut, dessen Vektor-DNA zusätzliche funktionelle Regulationssignale, beispielsweise Sequenzen für Replikation oder Integration enthält. Geeignete Vektoren sind unter anderem in "Methods in Plant Molecular Biology and Biotechnology" (CRC Press), Kap. 6/7, S. 71-119 (1993) beschrieben.

- 20 Unter Verwendung der oben zitierten Rekombinations- und Klonierungstechniken können die Expressionskassetten in geeignete Vektoren kloniert werden, die ihre Vermehrung, beispielsweise in *E. coli*, ermöglichen. Geeignete Klonierungsvektoren sind u.a. pJIT117 (Guerineau et al. (1988) Nucl. Acids Res. 16 :11380), pBR332, pUC-Serien, M13mp-Serien und pACYC184. Besonders geeignet sind binäre Vektoren, die sowohl in *E. coli* als auch in Agrobakterien replizieren können.

- 25 30 Im folgenden wird die Herstellung der erfindungsgemäßen genetisch veränderten Mikroorganismen näher beschrieben:

- Die vorstehend beschriebenen Nukleinsäuren, kodierend eine Ketolase oder  $\beta$ -Hydroxylase oder  $\beta$ -Cyclase sind vorzugsweise in Expressionskonstrukte eingebaut, enthaltend unter der genetischen Kontrolle regulativer Nukleinsäuresequenzen eine für ein erfindungsgemäßes Enzym kodierende Nukleinsäuresequenz; sowie Vektoren, umfassend wenigstens eines dieser Expressionskonstrukte.

Vorzugsweise umfassen solche erfindungsgemäßen Konstrukte 5'-stromaufwärts von der jeweiligen kodierenden Sequenz einen Promotor und 3'-stromabwärts eine Terminatorsequenz sowie gegebenenfalls weitere übliche regulative Elemente, und zwar jeweils operativ verknüpft mit der kodierenden Sequenz. Unter einer "operativen Verknüpfung" versteht man die sequentielle Anordnung von Promotor, kodierender Sequenz, Terminator und gegebenenfalls weiterer regulativer Elemente derart, dass jedes der regulativen Elemente seine Funktion bei der Expression der kodierenden Sequenz bestimmungsgemäß erfüllen kann.

Beispiele für operativ verknüpfbare Sequenzen sind Targeting-Sequenzen sowie Translationsverstärker, Enhancer, Polyadenylierungssignale und dergleichen. Weitere regulative Elemente umfassen selektierbare Marker, Amplifikationssignale, Replikationsursprünge und dergleichen.

Zusätzlich zu den artifiziellen Regulationssequenzen kann die natürliche Regulationssequenz vor dem eigentlichen Strukturgen noch vorhanden sein. Durch genetische Veränderung kann diese natürliche Regulation gegebenenfalls ausgeschaltet und die Expression der Gene erhöht oder erniedrigt werden. Das Genkonstrukt kann aber auch einfacher aufgebaut sein, das heißt es werden keine zusätzlichen Regulationssignale vor das Strukturgen insertiert und der natürliche Promotor mit seiner Regulation wird nicht entfernt. Statt dessen wird die natürliche Regulationssequenz so mutiert, dass keine Regulation mehr erfolgt und die Genexpression gesteigert oder verringert wird. Die Nukleinsäuresequenzen können in einer oder mehreren Kopien im Genkonstrukt enthalten sein.

Beispiele für brauchbare Promotoren in Mikroorganismen sind: cos-, tac-, trp-, tet-, trp-tet-, lpp-, lac-, lpp-lac-, lacIq-, T7-, T5-, T3-, gal-, trc-, ara-, SP6-, lambda-PR- oder im lambda-PL-Promotor, die vorteilhafterweise in gram-negativen Bakterien Anwendung finden; sowie die gram-positiven Promotoren amy und SPO2 oder die Hefepromotoren ADC1, MFa, AC, P-60, CYC1, GAPDH. Besonders bevorzugt ist die Verwendung induzierbarer Promotoren, wie z.B. licht- und insbesondere temperaturinduzierbarer Promotoren, wie der P<sub>1</sub>P<sub>1</sub>-Promotor.

Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen verwendet werden. Darüber hinaus können auch synthetische Promotoren vorteilhaft verwendet werden.

Die genannten regulatorischen Sequenzen sollen die gezielte Expression der Nukleinsäuresequenzen und die Proteinexpression ermöglichen. Dies kann beispielsweise je nach Wirtszell bedeuten, dass das Gen erst nach Induktion exprimiert oder überexprimiert wird, oder dass es sofort exprimiert und/oder überexprimiert wird.

Die regulatorischen Sequenzen bzw. Faktoren können dabei vorzugsweise die Expression positiv beeinflussen und dadurch erhöhen oder erniedrigen. So kann eine Verstärkung der regulato-



rischen Elemente vorteilhafterweise auf der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird.

5

Die Herstellung einer Expressionskassette erfolgt durch Fusion eines geeigneten Promotors mit den vorstehend beschriebenen Nukleinsäuresequenzen, kodierend eine Ketolase,  $\beta$ -

Hydroxylase oder  $\beta$ -Cyclase sowie einem Terminator- oder Polyadenylierungssignal. Dazu verwendet man gängige Rekombinations- und Klonierungstechniken, wie sie beispielsweise in T.

10

Maniatis, E.F. Fritsch und J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in T.J. Silhavy, M.L. Berman und L.W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel, F.M. et al., *Current Protocols in Molecular Biology*, Greene Publishing Assoc. and Wiley Interscience (1987) beschrieben sind.

15

Das rekombinante Nukleinsäurekonstrukt bzw. Genkonstrukt wird zur Expression in einem geeigneten Wirtsorganismus vorteilhafterweise in einen wirtsspezifischen Vektor inseriert, der eine optimale Expression der Gene im Wirt ermöglicht. Vektoren sind dem Fachmann wohl bekannt und können beispielsweise aus "Cloning Vectors" (Pouwels P. H. et al., Hrsg, Elsevier, Amsterdam-New York-Oxford, 1985) entnommen werden. Unter Vektoren sind außer Plasmiden auch alle anderen dem Fachmann bekannte Vektoren, wie beispielsweise Phagen, Viren, wie SV40, CMV, Baculovirus und Adenovirus, Transposons, IS-Elemente, Phasmide, Cosmide, und lineare oder zirkuläre DNA zu verstehen. Diese Vektoren können autonom im Wirtsorganismus repliziert oder chromosomal repliziert werden.

25

Als Beispiele für geeignete Expressionsvektoren können genannt werden:

30

Übliche Fusionsexpressionsvektoren, wie pGEX (Pharmacia Biotech Inc; Smith, D.B. und Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) und pRIT 5 (Pharmacia, Piscataway, NJ), bei denen Glutathion-S-Transferase (GST), Maltose E-bindendes Protein bzw. Protein A an das rekombinante Zielprotein fusioniert wird.

35

Nicht-Fusionsprotein-Expressionsvektoren wie pTrc (Amann et al., (1988) *Gene* 69:301-315) und pET 11d (Studier et al. *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Kalifornien (1990) 60-89) oder pBluescript und pUC-Vektoren.

Hefe-Expressionsvektor zur Expression in der Hefe *S. cerevisiae*, wie pYepSec1 (Baldari et al., (1987) *Embo J.* 6:229-234), pMFa (Kurjan und Herskowitz (1982) *Cell* 30:933-943), pJRY88

(Schultz et al. (1987) Gene 54:113-123) sowie pYES2 (Invitrogen Corporation, San Diego, CA).

Vektoren und Verfahren zur Konstruktion von Vektoren, die sich zur Verwendung in anderen Pilzen, wie filamentösen Pilzen, eignen, umfassen diejenigen, die eingehend beschrieben sind in: van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of Fungi, J.F. Peberdy et al., Hrsg., S. 1-28, Cambridge University Press: Cambridge.

Baculovirus-Vektoren, die zur Expression von Proteinen in gezüchteten Insektenzellen (bspw. Sf9-Zellen) verfügbar sind, umfassen die pAc-Reihe (Smith et al., (1983) Mol. Cell Biol. 3:2156-2165) und die pVL-Reihe (Lucklow und Summers (1989) Virology 170:31-39).

Weitere geeignete Expressionssysteme für prokaryontische und eukaryotische Zellen sind in Kapitel 16 und 17 von Sambrook, J., Fritsch, E.F. und Maniatis, T., Molecular cloning: A Laboratory Manual, 2. Auflage, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989 beschrieben.

Mit Hilfe der erfindungsgemäßen Expressionskonstrukte bzw. Vektoren sind genetisch veränderte Mikroorganismen herstellbar, welche beispielsweise mit wenigstens einem erfindungsgemäßen Vektor transformiert sind.

Vorteilhafterweise werden die oben beschriebenen erfindungsgemäßen rekombinanten Konstrukte in ein geeignetes Wirtssystem eingebracht und exprimiert. Dabei werden vorzugsweise dem Fachmann bekannte geläufige Klonierungs- und Transfektionsmethoden, wie beispielsweise Co-Präzipitation, Protoplastenfusion, Elektroporation, retrovirale Transfektion und dergleichen, verwendet, um die genannten Nukleinsäuren im jeweiligen Expressionssystem zur Expression zu bringen. Geeignete Systeme werden beispielsweise in Current Protocols in Molecular Biology, F. Ausubel et al., Hrsg., Wiley Interscience, New York 1997, beschrieben.

Die Selektion erfolgreich transformierter Organismen kann durch Markergene erfolgen, die ebenfalls im Vektor oder in der Expressionskassette enthalten sind. Beispiele für solche Markergene sind Gene für Antibiotikaresistenz und für Enzyme, die eine farbgebende Reaktion katalysieren, die ein Anfärben der transformierten Zelle bewirkt. Diese können dann mittels automatischer Zellsortierung selektiert werden.

Erfolgreich mit einem Vektor transformierte Mikroorganismen, die ein entsprechendes Antibiotikaresistenzgen (z.B. G418 oder Hygromycin) tragen, lassen sich durch entsprechende Antibiotika-enthaltende Medien oder Nährböden selektieren. Markerproteine, die an der Zelloberfläche

präsentiert werden, können zur Selektion mittels Affinitätschromatographie genutzt werden.

Die Kombination aus den Wirtsorganismen und den zu den Organismen passenden Vektoren, wie Plasmide, Viren oder Phagen, wie beispielsweise Plasmide mit dem RNA-Polymerase/Promoter-System, die Phagen 8 oder andere temperente Phagen oder Transposons und/oder weiteren vorteilhaften regulatorischen Sequenzen bildet ein Expressionssystem.

Die Erfindung betrifft ferner ein Verfahren zur Herstellung von genetisch veränderten Organismen, dadurch gekennzeichnet, das man ein Nukleinsäurekonstrukt, enthaltend funktionell verknüpft einen Promotor und Nukleinsäuren, kodierend eine Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist, und gegebenenfalls einen Terminator in das Genom des Ausgangsorganismus oder extrachromosomal in den Ausgangsorganismus einführt.

Die Erfindung betrifft ferner die genetisch veränderten Organismen, wobei die genetische Veränderung die Aktivität einer Ketolase

A für den Fall, dass der Wildtyporganismus bereits eine Ketolase-Aktivität aufweist, gegenüber dem Wildtyp erhöht und

B für den Fall, dass der Wildtyporganismus keine Ketolase-Aktivität aufweist, gegenüber dem Wildtyp verursacht

und die nach A erhöhte oder nach B verursachte Ketolase-Aktivität durch eine Ketolase verursacht wird, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.

Wie vorstehend ausgeführt erfolgt die Erhöhung oder Verursachung der Ketolase-Aktivität gegenüber dem Wildtyp vorzugsweise durch eine Erhöhung oder Verursachung der Genexpression einer Nukleinsäure, kodierend eine Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.

In einer weiter bevorzugten Ausführungsform erfolgt, wie vorstehend ausgeführt, die Erhöhung oder Verursachung der Genexpression einer Nukleinsäure, kodierend eine Ketolase, durch Ein-

bringen von Nukleinsäuren, kodierend eine Ketolase, in die Pflanzen und damit vorzugsweise durch Überexpression oder transgene Expression von Nukleinsäuren, kodierend eine Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist.

Die Erfindung betrifft ferner einen genetisch veränderten Organismus, enthaltend mindestens eine transgene Nukleinsäure, kodierend eine Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist. Dies ist der Fall, wenn der Ausgangsorganismus keine Ketolase oder eine endogen Ketolase aufweist und eine transgene Ketolase überexprimiert wird.

Die Erfindung betrifft ferner einen genetisch veränderten Organismus, enthaltend mindestens zwei endogene Nukleinsäuren, kodierend eine Ketolase, enthaltend die Aminosäuresequenz SEQ. ID. NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 42 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 2 aufweist. Dies ist der Fall, wenn der Ausgangsorganismus eine endogen Ketolase aufweist und die endogene Ketolase überexprimiert wird.

Besonders bevorzugte, genetisch veränderte Organismen weisen, wie vorstehend erwähnt, zusätzlich eine erhöhte Hydroxylase-Aktivität und/oder  $\beta$ -Cyclase-Aktivität gegenüber einem Wildtyporganismus auf. Weiter bevorzugte Ausführungsformen sind vorstehend im erfindungsgemäßen Verfahren beschrieben.

Unter Organismen werden erfindungsgemäß vorzugsweise Organismen verstanden, die als Wildtyp- oder Ausgangsorganismen natürlicherweise oder durch genetische Komplementierung und/oder Umregulierung der Stoffwechselwege in der Lage sind, Carotinoide, insbesondere  $\beta$ -Carotin und/oder Zeaxanthin und/oder Neoxanthin und/oder Violaxanthin und/oder Lutein herzustellen.

Weiter bevorzugte Organismen weisen als Wildtyp- oder Ausgangsorganismen bereits eine Hydroxylase-Aktivität auf und sind somit als Wildtyp- oder Ausgangsorganismen in der Lage, Zeaxanthin herzustellen.

Bevorzugte Organismen sind Pflanzen oder Mikroorganismen, wie beispielsweise Bakterien, Hefen, Algen oder Pilze.

Als Bakterien können sowohl Bakterien verwendet werden, die aufgrund des Einbringens von Genen der Carotinoidbiosynthese eines Carotinoid-produzierenden Organismus in der Lage sind, Xanthophylle zu synthetisieren, wie beispielsweise Bakterien der Gattung *Escherichia*, die beispielsweise crt-Gene aus *Erwinia* enthalten, als auch Bakterien, die von sich aus in der Lage sind, Xanthophylle zu synthetisieren wie beispielsweise Bakterien der Gattung *Erwinia*, *Agrobacterium*, *Flavobacterium*, *Alcaligenes*, *Paracoccus*, *Nostoc* oder Cyanobakterien der Gattung *Synechocystis*.

Bevorzugte Bakterien sind *Escherichia coli*, *Erwinia herbicola*, *Erwinia uredovora*, *Agrobacterium aurantiacum*, *Alcaligenes* sp. PC-1, *Flavobacterium* sp. strain R1534, das Cyanobacterium *Synechocystis* sp. PCC6803, *Paracoccus marcusii* oder *Paracoccus carotinifaciens*.

Bevorzugte Hefen sind *Candida*, *Saccharomyces*, *Hansenula*, *Pichia* oder *Phaffia*. Besonders bevorzugte Hefen sind *Xanthophyllomyces dendrorhous* oder *Phaffia rhodozyma*.

Bevorzugte Pilze sind *Aspergillus*, *Trichoderma*, *Ashbya*, *Neurospora*, *Blakeslea*, *Phycomyces*, *Fusarium* oder weitere in Indian Chem. Engr. Section B. Vol. 37, No. 1, 2 (1995) auf Seite 15, Tabelle 6 beschriebene Pilze.

Bevorzugte Algen sind Grünalgen, wie beispielsweise Algen der Gattung *Haematococcus*, *Phaedactylum tricornutum*, *Volvox* oder *Dunaliella*. Besonders bevorzugte Algen sind *Haematococcus puvialis* oder *Dunaliella bardawil*.

Weitere brauchbare Mikroorganismen und deren Herstellung zur Durchführung des erfindungsgemäßen Verfahrens sind beispielsweise aus der DE-A-199 16 140 bekannt, worauf hiermit Bezug genommen wird.

Besonders bevorzugte Pflanzen sind Pflanzen ausgewählt aus den Familien Ranunculaceae, Berberidaceae, Papaveraceae, Cannabaceae, Rosaceae, Fabaceae, Linaceae, Vitaceae, Brassicaceae, Cucurbitaceae, Primulaceae, Caryophyllaceae, Amaranthaceae, Gentianaceae, Geraniaceae, Caprifoliaceae, Oleaceae, Tropaeolaceae, Solanaceae, Scrophulariaceae, Astera-ceae, Liliaceae, Amaryllidaceae, Poaceae, Orchidaceae, Malvaceae, Iliaceae oder Lamiaceae.

Ganz besonders bevorzugte Pflanzen sind ausgewählt aus der Gruppe der Pflanzengattungen *Marigold*, *Tagetes erecta*, *Tagetes patula*, *Acacia*, *Aconitum*, *Adonis*, *Arnica*, *Aquilegia*, *Aster*, *Astragalus*, *Bignonia*, *Calendula*, *Caltha*, *Campanula*, *Canna*, *Centaurea*, *Cheiranthus*, *Chrysanthemum*, *Citrus*, *Crepis*, *Crocus*, *Curcubita*, *Cytisus*, *Delonia*, *Delphinium*, *Dianthus*, *Dimorphotheca*, *Doronicum*, *Eschscholtzia*, *Forsythia*, *Fremontia*, *Gazania*, *Gelsemium*, *Genista*, *Gentiana*, *Geranium*, *Gerbera*, *Geum*, *Grevillea*, *Helenium*, *Helianthus*, *Hepatica*, *Heracleum*,

*Hisbiscus, Heliopsis, Hypericum, Hypochoeris, Impatiens, Iris, Jacaranda, Kerria, Laburnum, Lathyrus, Leontodon, Lilium, Linum, Lotus, Lycopersicon, Lysimachia, Marattia, Medicago, Mimulus, Narcissus, Oenothera, Osmanthus, Petunia, Photinia, Physalis, Phyteuma, Potentilla, Pyracantha, Ranunculus, Rhododendron, Rosa, Rudbeckia, Senecio, Silene, Silphium, Sinapsis, Sorbus, Spartium, Tecoma, Torenia, Tragopogon, Trollius, Tropaeolum, Tulipa, Tussilago, Ulex, Viola oder Zinnia*, besonders bevorzugt ausgewählt aus der Gruppe der Pflanzengattungen Marigold, Tagetes erecta, Tagetes patula, Lycopersicon, Rosa, Calendula, Physalis, Medicago, Helianthus, Chrysanthemum, Aster, Tulipa, Narcissus, Petunia, Geranium, Tropaeolum oder Adonis.

10

Ganz besonders bevorzugte genetisch veränderte Pflanzen sind ausgewählt aus den Pflanzengattungen Marigold, Tagetes erecta, Tagetes patula, Adonis, Lycopersicon, Rosa, Calendula, Physalis, Medicago, Helianthus, Chrysanthemum, Aster, Tulipa, Narcissus, Petunia, Geranium oder Tropaeolum, wobei die genetisch veränderte Pflanze mindestens eine transgene Nukleinsäure, kodierend eine Ketolase, enthält.

15

Die transgenen Pflanzen, deren Vermehrungsgut, sowie deren Pflanzenzellen, -gewebe oder -teile, insbesondere deren Früchte, Samen, Blüten und Blütenblätter sind ein weiterer Gegenstand der vorliegenden Erfindung.

20

Die genetisch veränderten Pflanzen können, wie vorstehend beschrieben, zur Herstellung von Ketocarotinoiden, insbesondere Astaxanthin verwendet werden.

25

Von Menschen und Tieren verzehrbare erfindungsgemäße, genetisch veränderte Organismen, insbesondere Pflanzen oder Pflanzenteile, wie insbesondere Blütenblätter mit erhöhtem Gehalt an Ketocarotinoiden, insbesondere Astaxanthin können auch beispielsweise direkt oder nach an sich bekannter Prozessierung als Nahrungsmittel oder Futtermittel oder als Futter- und Nahrungsergänzungsmittel verwendet werden.

30

Ferner können die genetisch veränderten Organismen zur Herstellung von Ketocarotinoidhaltigen Extrakten der Organismen und/oder zur Herstellung von Futter- und Nahrungsergänzungsmitteln verwendet werden.

35

Die genetisch veränderten Organismen weisen im Vergleich zum Wildtyp einen erhöhten Gehalt an Ketocarotinoiden auf.

Unter einem erhöhten Gehalt an Ketocarotinoiden wird in der Regel ein erhöhter Gehalt an Gesamt-Ketocarotinoid verstanden.

Unter einem erhöhten Gehalt an Ketocarotinoiden wird aber auch insbesondere ein veränderter Gehalt der bevorzugten Ketocarotinoide verstanden, ohne dass zwangsläufig der Gesamt-Carotinoidgehalt erhöht sein muss.

- 5 In einer besonders bevorzugten Ausführungsform weisen die erfindungsgemäßen, genetisch veränderten Pflanzen im Vergleich zum Wildtyp einen erhöhten Gehalt an Astaxanthin auf.

Unter einem erhöhten Gehalt wird in diesem Fall auch ein verursachter Gehalt an Ketocarotinoiden, bzw. Astaxanthin verstanden.

10

Die Erfindung betrifft ferner die neuen Ketolasen sowie die neuen Nukleinsäuren, die diese kodieren.

- 15 Insbesondere betrifft die Erfindung Ketolasen, enthaltend die Aminosäuresequenz SEQ. ID. NO. 8 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 70 %, vorzugsweise mindestens 75%, besonders bevorzugt mindestens 80%, bevorzugter mindestens 85%, bevorzugter mindestens 90%, bevorzugter mindestens 95% auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 8 aufweist, mit der Maßgabe, dass die Aminosäuresequenzen SEQ ID NO: 4 nicht enthalten ist. Die Sequenz SEQ ID NO: 4 ist, wie vorstehend erwähnt, als putatives Protein in Datenbanken annotiert.

- 25 Ferner betrifft die Erfindung Ketolasen, enthaltend die Aminosäuresequenz SEQ. ID. NO. 6 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 70 % auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 6 aufweist. Die Sequenz SEQ ID NO: 6 ist, wie vorstehend erwähnt, in Datenbanken nicht annotiert.

- 30 In einer weiteren Ausführungsform betrifft die Erfindung Ketolasen, enthaltend die Aminosäuresequenz SEQ. ID. NO. 12 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 70 %, vorzugsweise mindestens 75%, besonders bevorzugt mindestens 80%, bevorzugter mindestens 85%, bevorzugter mindestens 90%, bevorzugter mindestens 95% auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 12 aufweist, mit der Maßgabe, dass die Aminosäuresequenzen SEQ ID NO: 6 nicht enthalten ist.

35

Ferner betrifft die Erfindung Ketolasen, enthaltend die Aminosäuresequenz SEQ. ID. NO. 49 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 50 %, vorzugsweise mindestens 60%, be-

sonders bevorzugt mindestens 70%, bevorzugter mindestens 80%, bevorzugter mindestens 90%, bevorzugter mindestens 95% auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 49 aufweist, mit der Maßgabe, dass die Aminosäuresequenzen SEQ ID NO: 47 nicht enthalten ist. Die Sequenz SEQ ID NO: 47 ist, wie vorstehend erwähnt, als putatives Protein in Datenbanken

5 annotiert.

Die Erfindung betrifft ferner Nukleinsäuren, kodierend ein vorstehend beschriebenes Protein, mit der Maßgabe, dass die Nukleinsäure nicht die Sequenz SEQ ID NO: 5 enthält.

10 Überraschenderweise wurde gefunden, dass ein Protein enthaltend die Aminosäuresequenz SEQ. ID. NO. 4 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 70 %, vorzugsweise mindestens 75%, besonders bevorzugt mindestens 80%, bevorzugter mindestens 85%, bevorzugter mindestens 90%, bevorzugter mindestens 95% auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 4 und die Eigenschaft einer Ketolase aufweist, eine Eigenschaft als Ketolase aufweist.

15 Die Erfindung betrifft daher auch die Verwendung eines Proteins, enthaltend die Aminosäuresequenz SEQ. ID. NO. 4 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 70 %, vorzugsweise mindestens 75%, besonders bevorzugt mindestens 80%, bevorzugter mindestens 85%, bevorzugter mindestens 90%, bevorzugter mindestens 95% auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 4 und die Eigenschaft einer Ketolase aufweist, als Ketolase.

20 Ferner wurde überraschenderweise gefunden, dass ein Protein enthaltend die Aminosäuresequenz SEQ. ID. NO. 6 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 65%, vorzugsweise mindestens 70 %, vorzugsweise mindestens 75%, besonders bevorzugt mindestens 80%, bevorzugter mindestens 85%, bevorzugter mindestens 90%, bevorzugter mindestens 95% auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 6 und die Eigenschaft einer Ketolase aufweist, eine Eigenschaft als Ketolase aufweist.

25 Die Erfindung betrifft daher auch die Verwendung eines Proteins, enthaltend die Aminosäuresequenz SEQ. ID. NO. 6 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 65%, vorzugsweise mindestens 70 %, vorzugsweise mindestens 75%, besonders bevorzugt mindestens 80%, bevorzugter mindestens 85%, bevorzugter mindestens 90%, bevorzugter mindestens 95% auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 6 und die Eigenschaft einer Ketolase aufweist, als Ketolase.



Ferner wurde überraschenderweise gefunden, dass ein Protein enthaltend die Aminosäuresequenz SEQ. ID. NO. 47 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 50%, vorzugsweise mindestens 60 %, vorzugsweise mindestens 70%, besonders bevorzugt mindestens 80%, bevorzugter mindestens 85%, bevorzugter mindestens 90%, bevorzugter mindestens 95% auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 47 und die Eigenschaft einer Ketolase aufweist, eine Eigenschaft als Ketolase aufweist.

Die Erfindung betrifft daher auch die Verwendung eines Proteins, enthaltend die Aminosäuresequenz SEQ. ID. NO. 47 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 50%, vorzugsweise mindestens 60 %, vorzugsweise mindestens 70%, besonders bevorzugt mindestens 80%, bevorzugter mindestens 85%, bevorzugter mindestens 90%, bevorzugter mindestens 95% auf Aminosäureebene mit der Sequenz SEQ. ID. NO. 47 und die Eigenschaft einer Ketolase aufweist, als Ketolase.

Im Vergleich zu den Verfahren des Standes der Technik, liefert das erfindungsgemäße Verfahren eine höhere Menge an Ketocarotinoide, insbesondere Astaxanthin mit einer geringeren Menge an hydroxylierten Nebenprodukten.

Die Erfindung wird durch die nun folgenden Beispiele erläutert, ist aber nicht auf diese beschränkt:

Allgemeine Experimentelle Bedingungen:  
Sequenzanalyse rekombinanter DNA

Die Sequenzierung rekombinanter DNA-Moleküle erfolgte mit einem Laserfluoreszenz-DNA-Sequenzierer der Firma Licor (Vertrieb durch MWG Biotech, Ebersbach) nach der Methode von Sanger (Sanger et al., Proc. Natl. Acad. Sci. USA 74 (1977), 5463-5467).

Beispiel 1:

Amplifikation einer DNA, die die gesamte Primärsequenz der Ketolase aus *Nostoc sp. PCC 7120* codiert

Die DNA, die für die Ketolase aus *Nostoc PCC 7120* kodiert, wurde mittels PCR aus *Nostoc PCC 7120* (Stamm der "Pasteur Culture Collection of Cyanobacterium") amplifiziert.

Für die Präparation von genomischer DNA aus einer Suspensionskultur von *Nostoc PCC 7120*, die 1 Woche mit Dauerlicht und konstantem Schütteln (150 rpm) at 25°C in BG 11-Medium (1.5

g/l NaNO<sub>3</sub>, 0.04 g/l K<sub>2</sub>PO<sub>4</sub>·3H<sub>2</sub>O, 0.075 g/l MgSO<sub>4</sub>·xH<sub>2</sub>O, 0.036 g/l CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.006 g/l citric acid, 0.006 g/l Ferric ammonium citrate, 0.001 g/l EDTA disodium magnesium, 0.04 g/l Na<sub>2</sub>CO<sub>3</sub>, 1ml trace metal mix A5+Co (2.86 g/l H<sub>3</sub>BO<sub>3</sub>, 1.81 g/l MnCl<sub>2</sub>·4H<sub>2</sub>O, 0.222 g/l ZnSO<sub>4</sub>·7H<sub>2</sub>O, 0.39 g/l NaMoO<sub>4</sub>·2H<sub>2</sub>O, 0.079 g/l CuSO<sub>4</sub>·5H<sub>2</sub>O, 0.0494 g/l Co(NO<sub>3</sub>)<sub>2</sub>·6H<sub>2</sub>O)

- 5 gewachsen war, wurden die Zellen durch Zentrifugation geerntet, in flüssigem Stickstoff eingefroren und im Mörser pulverisiert.

#### Protokoll für DNA Isolation aus *Nostoc* PCC7120:

- 10 Aus einer 10 ml Flüssigkultur wurden die Bakterienzellen durch 10minütige Zentrifugation bei 8 000 rpm pelletiert. Anschließend wurden die Bakterienzellen in flüssigem Stickstoff mit einem Mörser zerstoßen und gemahlen. Das Zellmaterial wurde in 1 ml 10mM Tris HCl (pH 7.5) re-
- 15 suspendiert und in ein Eppendorf Reaktionsgefäß (2ml Volumen) überführt. Nach Zugabe von 100 µl Proteinase K (Konzentration: 20 mg/ml) wurde die Zellsuspension für 3 Stunden bei 37°C inkubiert. Anschließend wurde die Suspension mit 500 µl Phenol extrahiert. Nach 5minütiger
- Zentrifugation bei 13 000 upm wurde die obere, wässrige Phase in ein neues 2 ml-Eppendorf Reaktionsgefäß überführt. Die Extraktion mit Phenol wurde 3mal wiederholt. Die DNA wurde durch Zugabe von 1/10 Volumen 3 M Natriumacetat (pH 5.2) und 0.6 Volumen Isopropanol ge-
- 20 fällt und anschließend mit 70% Ethanol gewaschen. Das DNA-Pellet wurde bei Raumtemperatur getrocknet, in 25 µl Wasser aufgenommen und unter Erhitzung auf 65°C gelöst.

- Die Nukleinsäure, kodierend eine Ketolase aus *Nostoc* PCC 7120, wurde mittels "polymerase chain reaction" (PCR) aus *Nostoc* PCC 7120 unter Verwendung eines sense-spezifischen Primers (NOSTF, SEQ ID No. 19) und eines antisense-spezifischen Primers (NOSTG SEQ ID No.
- 25 20) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

- Die PCR zur Amplifikation der DNA, die für ein Ketolase Protein bestehend aus der gesamten
- 30 Primärsequenz kodiert, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

- 1 µl einer *Nostoc* PCC 7120 DNA (hergestellt wie oben beschrieben)
- 0.25 mM dNTPs
- 0.2 mM NOSTF (SEQ ID No. 19)
- 35 - 0.2 mM NOSTG (SEQ ID No. 20)
- 5 µl 10X PCR-Puffer (TAKARA)
- 0.25 µl R Taq Polymerase (TAKARA)
- 25.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X	94°C	2 Minuten
35X	94°C	1 Minute
5	55°C	1 Minuten
	72°C	3 Minuten
1X	72°C	10 Minuten

Die PCR-Amplifikation mit SEQ ID No. 19 und SEQ ID No. 20 resultierte in einem 805 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (SEQ ID No. 21). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pGEM-T (Promega) kloniert und der Klon pNOSTF-G erhalten.

Sequenzierung des Klons pNOSTF-G mit dem M13F- und dem M13R-Primer bestätigte eine Sequenz, welche mit der DNA-Sequenz des Datenbankeintrages AP003592 identisch ist. Diese Nukleotidsequenz wurde in einem unabhängigen Amplifikationsexperiment reproduziert und repräsentiert somit die Nukleotidsequenz im verwendeten *Nostoc PCC 7120*.

Dieser Klon pNOSTF-G wurde daher für die Klonierung in den Expressionsvektor pJIT117 (Guérineau et al. 1988, Nucl. Acids Res. 16: 11380) verwendet. Die Klonierung erfolgte durch Isolierung des 1027 Bp SphI-Fragmentes aus pGEM-T und Ligierung in den SphI geschnittenen Vektor pJIT117. Der Klon, der die Ketolase von *Nostoc* in der korrekten Orientierung als N-terminale translationale Fusion mit dem rbcS Transitpeptid enthält, heisst pJNOST.

#### Beispiel 2:

Konstruktion des Plasmides pMCL-CrtYIBZ/idi/gps für die Synthese von Zeaxanthin in *E. coli*

Die Konstruktion von pMCL-CrtYIBZ/idi/gps erfolgte in drei Schritten über die Zwischenstufen pMCL-CrtYIBZ und pMCL-CrtYIBZ/idi. Als Vektor wurde das mit high-copy-number Vektoren kompatible Plasmid pMCL200 verwendet (Nakano, Y., Yoshida, Y., Yamashita, Y. und Koga, T.; Construction of a series of pACYC-derived plasmid vectors; Gene 162 (1995), 157-158).

#### Beispiel 2.1.: Konstruktion von pMCL-CrtYIBZ

Die Biosynthesegene *crtY*, *crtB*, *crtI* und *crtZ* entstammen dem Bakterium *Erwinia uredovora* und wurden mittels PCR amplifiziert. Genomische DNA von *Erwinia uredovora* (DSM 30080) wurde von der Deutschen Sammlung von Mikroorganismen und Zellkulturen (DSMZ, Braunschweig) innerhalb eines Service-Dienstes präpariert. Die PCR-Reaktion wurde entsprechend den Angaben des Herstellers durchgeführt (Roche, Long Template PCR: Procedure for

amplification of 5-20 kb targets with the expand long template PCR system). Die PCR-Bedingungen für die Amplifikation des Biosynthesecusters von *Erwinia uredovora* waren die folgenden:

5 Master Mix 1:

- 1.75 µl dNTPs (Endkonzentration 350 µM)
  - 0.3 µM Primer Crt1 (SEQ ID No. 22)
  - 0.3 µM Primer Crt2 (SEQ ID No. 23)
  - 10 - 250 – 500 ng genomische DNA von DSM 30080
- Aq. Dest. bis zu einem Gesamtvolumen von 50 µl

Master Mix 2:

- 15 - 5 µl 10x PCR Puffer 1 (Endkonzentration 1x, mit 1.75 mM Mg<sup>2+</sup>)
  - 10x PCR Puffer 2 (Endkonzentration 1x, mit 2.25 mM Mg<sup>2+</sup>)
  - 10x PCR Puffer 3 (Endkonzentration 1x, mit 2.25 mM Mg<sup>2+</sup>)
  - 0.75 µl Expand Long Template Enzyme Mix (Endkonzentration 2.6 Units)
- Aq. Dest. bis zu einem Gesamtvolumen von 50 µl

20

Die beiden Ansätze "Master Mix 1" und "Master Mix 2" wurden zusammenpipetiert. Die PCR wurde in einem Gesamtvolumen von 50 µl unter folgenden Zyklusbedingungen durchgeführt:

- |    |     |      |             |
|----|-----|------|-------------|
| 25 | 1X  | 94°C | 2 Minuten   |
|    | 30X | 94°C | 30 Sekunden |
|    |     | 58°C | 1 Minute    |
|    |     | 68°C | 4 Minuten   |
|    | 1X  | 72°C | 10 Minuten  |

30

Die PCR-Amplifikation mit SEQ ID No. 22 und SEQ ID No. 23 resultierte in einem Fragment (SEQ ID NO: 24), das für die Gene *CrtY* (Protein: SEQ ID NO: 25), *CrtI* (Protein: SEQ ID NO: 26), *crtB* (Protein: SEQ ID NO: 27) und *CrtZ* (*idNA*) kodiert. Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pCR2.1 (Invitrogen) kloniert und

35 der Klon pCR2.1-CrtYIBZ erhalten.

Das Plasmid pCR2.1-CrtYIBZ wurde Sall und HindIII geschnitten, das resultierende Sall/HindIII-Fragment isoliert und durch Ligierung in den Sall/HindIII geschnittenen Vektor pMCL200 transfe-

riert. Das in pMCL 200 klonierte Sall/HindIII Fragment aus pCR2.1-CrtYIBZ ist 4624 Bp lang, kodiert für die Gene *CrtY*, *CrtI*, *crtB* und *CrtZ* und entspricht der Sequenz von Position 2295 bis 6918 in D90087 (SEQ ID No. 24). Der resultierende Klon heisst pMCL-CrtYIBZ.

5 Beispiel 2.2.: Konstruktion von pMCL-CrtYIBZ/*idi*

Das Gen *idi* (Isopentenylidiphosphat-Isomerase; IPP-Isomerase) wurde aus *E. coli* mittels PCR amplifiziert. Die Nukleinsäure, kodierend das gesamte *idi* Gen mit *idi*-Promotor und Ribosomenbindestelle, wurde aus *E. coli* mittels "polymerase chain reaction" (PCR) unter Verwendung eines sense-spezifischen Primers (5'-*idi* SEQ ID No. 28) und eines antisense-spezifischen Primers (3'-*idi* SEQ ID No. 29) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der DNA erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

- 1 µl einer *E. coli* TOP10- Suspension
- 0.25 mM dNTPs
- 0.2 mM 5'-*idi* (SEQ ID No. 28)
- 0.2 mM 3'-*idi* (SEQ ID No. 29)
- 5 µl 10X PCR-Puffer (TAKARA)
- 0.25 µl R Taq Polymerase (TAKARA)
- 28.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X	94°C	2 Minuten
20X	94°C	1 Minute
	62 °C	1 Minute
	72°C	1 Minute
1X	72°C	10 Minuten

Die PCR-Amplifikation mit SEQ ID No. 28 und SEQ ID No. 29 resultierte in einem 679 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (SEQ ID No. 30). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pCR2.1 (Invitrogen) kloniert und der Klon pCR2.1-*idi* erhalten.

Sequenzierung des Klons pCR2.1-idi bestätigte eine Sequenz, die sich nicht von der publizierten Sequenz AE000372 in Position 8774 bis Position 9440 unterscheidet. Diese Region umfaßt die Promotor-Region, die potentielle Ribosomenbindestelle und den gesamten "open reading frame" für die IPP-Isomerase. Das in pCR2.1-idi klonierte Fragment hat durch das Einfügen einer XhoI-Schnittstelle am 5'-Ende und einer Sall-Schnittstelle am 3'-Ende des *idi*-Gens eine Gesamtlänge von 679 Bp.

Dieser Klon wurde daher für die Klonierung des *idi*-Gens in den Vektor pMCL-CrtYIBZ verwendet. Die Klonierung erfolgte durch Isolierung des XhoI/Sall-Fragmentes aus pCR2.1-idi und Ligierung in den XhoI/Sall geschnittenen Vektor pMCL-CrtYIBZ. Der resultierende Klon heisst pMCL-CrtYIBZ/idi.

#### Beispiel 2.3.: Konstruktion von pMCL-CrtYIBZ/idi/gps

Das Gen *gps* (Geranylgeranylpyrophosphat-Synthase; ; GGPP-Synthase) wurde aus *Archaeoglobus fulgidus* mittels PCR amplifiziert. Die Nukleinsäure, kodierend *gps* aus *Archaeoglobus fulgidus*, wurde mittels "polymerase chain reaction" (PCR) unter Verwendung eines sense-spezifischen Primers (5'-*gps* SEQ ID No. 32) und eines antisense-spezifischen Primers (3'-*gps* SEQ ID No. 33) amplifiziert.

Die DNA von *Archaeoglobus fulgidus* wurde von der Deutschen Sammlung von Mikroorganismen und Zellkulturen (DSMZ, Braunschweig) innerhalb eines Service-Dienstes präpariert. Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der DNA, die für ein GGPP-Synthase Protein bestehend aus der gesamten Primärsequenz kodiert, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

- 1 ul einer *Archaeoglobus fulgidus*-DNA
- 0.25 mM dNTPs
- 0.2 mM 5'-*gps* (SEQ ID No. 32)
- 0.2 mM 3'-*gps* (SEQ ID No. 33)
- 5 ul 10X PCR-Puffer (TAKARA)
- 0.25 ul R Taq Polymerase (TAKARA)
- 28.8 ul Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X	94°C	2 Minuten
20X	94°C	1 Minute

56°C 1 Minute

72°C 1 Minute

1X 72°C 10 Minuten

- 5 Das mittels PCR und den Primern SEQ ID No. 32 und SEQ ID No. 33 amplifizierte DNA-Fragment wurde mit an sich bekannten Methoden aus dem Agarosegel eluiert und mit den Restriktionsenzymen NcoI und HindIII geschnitten. Daraus resultiert ein 962 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (SEQ ID No. 34). Unter Verwendung von Standardmethoden wurde das NcoI/HindIII geschnittene Amplifikat in den Vektor pCB97-30 kloniert und der Klon pCB-gps erhalten.

- 15 Sequenzierung des Klons pCB-gps bestätigte eine Sequenz für die GGPP-Synthase aus *A. fulgidus*, die sich von der publizierten Sequenz AF120272 in einem Nukleotid unterscheidet. Durch das Einfügen einer NcoI-Schnittstelle im *gps*-Gen wurde das zweite Kodon der GGPP-Synthase verändert. In der publizierten Sequenz AF120272 kodiert CTG (Position 4-6) für Leucin. Durch die Amplifikation mit den beiden Primern SEQ ID No. 32 und SEQ ID No. 33 wurde dieses zweite Kodon in GTG verändert, welches für Valin kodiert.

- 20 Der Klon pCB-gps wurde daher für die Klonierung des *gps*-Gens in den Vektor pMCL-CrtYIBZ/idi verwendet. Die Klonierung erfolgte durch Isolierung des KpnI/XhoI-Fragmentes aus pCB-gps und Ligierung in den KpnI und XhoI geschnittenen Vektor pMCL-CrtYIBZ/idi. Das klonierte KpnI/XhoI-Fragment (SEQ ID No. 34) trägt den Prm16-Promotor zusammen mit einer minimalen 5'-UTR-Sequenz von *rbcl*, den ersten 6 Kodons von *rbcl*, die die GGPP-Synthase N-terminal verlängern, und 3' vom *gps*-Gen die *psbA*-Sequenz. Der N-Terminus der GGPP-Synthase hat somit anstelle der natürlichen Aminosäure-Abfolge mit Met-Leu-Lys-Glu (Aminosäure 1 bis 4 aus AF120272) die veränderte Aminosäure-Abfolge Met-Thr-Pro-Gln-Thr-Ala-Met-Val-Lys-Glu. Daraus resultiert, dass die rekombinante GGPP-Synthase, beginnend mit Lys in Position 3 (in AF120272) identisch ist und keine weiteren Änderungen in der Aminosäuresequenz aufweist. Die *rbcl*- und *psbA*-Sequenzen wurden gemäß einer Referenz nach Eibl et al. (Plant J. 19. (1999), 1-13) verwendet. Der resultierende Klon heisst pMCL-CrtYIBZ/idi/gps.

#### Beispiel 3:

Biotransformation von Zeaxanthin in rekombinanten *E. coli*-Stämmen

- 35 Zur Zeaxanthin-Biotransformation wurden rekombinante *E. coli*-Stämme hergestellt, welche durch heterologe Komplementation zur Zeaxanthin-Produktion befähigt sind. Stämme von *E. coli* TOP10 wurden als Wirtszellen für die Komplementations-Experimente mit den Plasmiden pNOSTF-G und pMCL-CrtYIBZ/idi/gps verwendet.

Um *E. coli*-Stämme herzustellen, die die Synthese von Zeaxanthin in hoher Konzentration ermöglichen, wurde das Plasmid pMCL-CrtYIBZ/idi/gps konstruiert. Das Plasmid trägt die Biosynthesegene *crtY*, *crtB*, *crtI* und *crtY* von *Erwinia uredovora*, das Gen *gps* (für Geranylgeranylpyrophosphat-Synthase) aus *Archaeoglobus fulgidus* und das Gen *idi* (Isopentenylidiphosphat-Isomerase) aus *E. coli*. Mit diesem Konstrukt wurden limitierende Schritte für eine hohe Akkumulation von Carotinoiden und deren biosynthetischen Vorstufen beseitigt. Dies wurde zuvor von Wang et al. in ähnlicher Weise mit mehreren Plasmiden beschrieben (Wang, C.-W., Oh, M.-K. und Liao, J.C.; Engineered isoprenoid pathway enhances astaxanthin production in Escherichia coli, Biotechnology and Bioengineering 62 (1999), 235-241).

Kulturen von *E. coli* TOP10 wurden in an sich bekannter Weise mit den beiden Plasmiden pNOSTF-G und pMCL-CrtYIBZ/idi/gps transformiert und in LB-Medium bei 30°C bzw. 37°C über Nacht kultiviert. Ampicillin (50 µg/ml), Chloramphenicol (50 µg/ml) und Isopropyl-β-thiogalactosid (1 mmol) wurden in an sich üblicher Weise ebenfalls über Nacht zugegeben.

Zur Isolierung der Carotinoide aus den rekombinanten Stämmen wurden die Zellen mit Aceton extrahiert, das organische Lösungsmittel zur Trockne eingedampft und die Carotinoide mittels HPLC über eine C30-Säule aufgetrennt. Folgende Verfahrensbedingungen wurden eingestellt.

Trennsäule: ProntoSil C30-Säule, 250 x 4,6 mm, (Bischoff, Leonberg)

Flussrate: 1.0 ml/min

Eluenten: Laufmittel A - 100% Methanol

Laufmittel B - 80% Methanol, 0.2% Ammoniumacetat

Laufmittel C - 100% t-Butyl-methylether

Gradientprofil:

Zeit	Flussrate	% Laufmittel A	% Laufmittel B	% Laufmittel C
1.00	1.0	95.0	5.0	0
1.05	1.0	80.0	5.0	15.0
14.00	1.0	42.0	5.0	53.0
14.05	1.0	95.0	5.0	0
17.00	1.0	95.0	5.0	0
18.00	1.0	95.0	5.0	0

Detektion: 300 - 500 nm



Die Spektren wurden direkt aus den Elutionspeaks unter Verwendung eines Photodiodenarray-detektors bestimmt. Die isolierten Substanzen wurden über ihre Absorptionsspektren und ihre Retentionszeiten im Vergleich zu Standardproben identifiziert.

- 5 Abbildung 1 zeigt die chromatographische Analyse einer Probe erhalten aus einem mit pNOSTF-G und pMCL-CrtYIBZ/idi/gps transformierten *E. coli*-Stamm. Es zeigt sich, daß dieser Stamm aufgrund der heterologen Komplementation verschiedene Ketocarotinoide synthetisieren kann. Mit zunehmender Retentionszeit werden Astaxanthin (Peak 1), Adonirubin (Peak 2) und Canthaxanthin (Peak 3) eluiert.

10

Beispiel 3.1

Vergleichsbeispiel

15

Analog zu den vorhergehenden Beispielen wurde als Vergleichsbeispiel ein *E.coli*-Stamm hergestellt, der eine Ketolase aus *Haematococcus pluvialis* Flotow em. Wille exprimiert. Dazu wurde die cDNA die für die gesamte Primärsequenz der Ketolase aus *Haematococcus pluvialis* Flotow em. Wille codiert amplifiziert und gemäß Beispiel 1 in den gleichen Expressionsvektor kloniert.

20

Die cDNA, die für die Ketolase aus *Haematococcus pluvialis* codiert, wurde mittels PCR aus *Haematococcus pluvialis* (Stamm 192.80 der "Sammlung von Algenkulturen der Universität Göttingen") Suspensionskultur amplifiziert. Für die Präparation von Total-RNA aus einer Suspensionskultur von *Haematococcus pluvialis* (Stamm 192.80), die 2 Wochen mit indirektem Tageslicht bei Raumtemperatur in *Haematococcus*- Medium (1.2 g/l Natriumacetat, 2 g/l Hefeextrakt, 0.2 g/l  $MgCl_2 \cdot 6H_2O$ , 0.02  $CaCl_2 \cdot 2H_2O$ ; pH 6.8; nach Autoklavieren Zugabe von 400 mg/l L-Asparagin, 10 mg/l  $FeSO_4 \cdot xH_2O$ ) gewachsen war, wurden die Zellen geerntet, in flüssigem Stickstoff eingefroren und im Mörser pulverisiert. Anschließend wurden 100 mg der gefrorenen, pulverisierten Algenzellen in ein Reaktionsgefäß überführt und in 0.8 ml Trizol-Puffer (Life Technologies) aufgenommen. Die Suspension wurde mit 0.2 ml Chloroform extrahiert.

25

- 30 Nach 15 minütiger Zentrifugation bei 12 000 g wurde der wässrige Überstand abgenommen und in ein neues Reaktionsgefäß überführt und mit einem Volumen Ethanol extrahiert. Die RNA wurde mit einem Volumen Isopropanol gefällt, mit 75% Ethanol gewaschen und das Pellet in DEPC Wasser (über Nacht Inkubation von Wasser mit 1/1000 Volumen Diethylpyrocarbonat bei Raumtemperatur, anschließend autoklaviert) gelöst. Die RNA-Konzentration wurde photometrisch bestimmt.
- 35

Für die cDNA-Synthese wurden 2.5 µg Gesamt-RNA für 10 min bei 60 °C denaturiert, für 2 min auf Eis abgekühlt und mittels eines cDNA-Kits (Ready-to-go-you-prime-beads, Pharmacia Biotech) nach Herstellerangaben unter Verwendung eines antisense spezifischen Primers PR1

(gcaagctcga cagctacaaa cc) in cDNA umgeschrieben.

- Die Nukleinsäure codierend eine Ketolase aus *Haematococcus pluvialis* (Stamm 192.80) wurde mittels polymerase chain reaction (PCR) aus *Haematococcus pluvialis* unter Verwendung eines sense spezifischen Primers PR2 (gaagcatgca gctagcagcg acag) und eines antisense spezifischen Primers PR1 amplifiziert.

Die PCR-Bedingungen waren die folgenden:

- Die PCR zur Amplifikation der cDNA, die für ein Ketolase Protein bestehend aus der gesamten Primärsequenz codiert, erfolgte in einem 50 ml Reaktionsansatz, in dem enthalten war:

- 4 ml einer *Haematococcus pluvialis* cDNA (hergestellt wie oben beschrieben)
- 0.25 mM dNTPs
- 0.2 mM PR1
- 0.2 mM PR2
- 5 ml 10X PCR-Puffer (TAKARA)
- 0.25 ml R Taq Polymerase (TAKARA)
- 25.8 ml Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X	94_C	2 Minuten
35X	94_C	1 Minute
25	53_C	2 Minuten
	72_C	3 Minuten
1X	72_C	10 Minuten

Die PCR-Amplifikation mit PR1 und PR2 resultierte in einem 1155 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz codiert:

	gaagcatgca	gctagcagcg	acagtaatgt	tggagcagct	taccggaagc	gctgaggcac	60
	tcaaggagaa	ggagaaggag	gttgacaggca	gctctgacgt	gttgcgtaga	tgggagaccc	120
	agtactcgct	tccgtcagag	gagtcagacg	cggcccgcgc	gggactgaag	aatgcctaca	180
35	agccaccacc	ttccgacaca	aagggcatca	caatggcgct	agctgtcatc	ggctcctggg	240
	ccgcagtggt	cctccacgcc	atttttcaaa	tcaagcttcc	gacctccttg	gaccagctgc	300
	actggctgcc	cgtgtcagat	gccacagctc	agctggtag	cggcagcagc	agcctgctgc	360
	acatcgtcgt	agtattcttt	gtcctggagt	tcctgtacac	aggccttttt	atcaccacgc	420
	atgatgctat	gcatggcacc	atcgccatga	gaaacaggca	gcttaatgac	ttcttgggca	480
40	gagtatgcat	ctccttgtac	gcctggtttg	attacaacat	gctgcaccgc	aagcattggg	540
	agcaccacaa	ccacactggc	gaggtgggca	aggacctga	cttccacagg	ggaaaccctg	600
	gcattgtgcc	ctggtttgcc	agcttcagtg	ccagctacat	gtcgatgtgg	cagtttgogc	660
	gcctcgcatg	gtggacggcg	gtcatgcagc	tgctgggtgc	gccaatggcg	aacctgctgg	720
	gtttcatggc	ggccgcgccc	atcctgtccg	ccttcgcgtt	gttctacttt	ggcacgtaca	780
45	tgccccacaa	gcctgagcct	ggcgccgcgt	caggctcttc	accagccgtc	atgaactggg	840
	ggaagtgcgc	cactagccag	gcgtccgacc	tggtcagctt	tctgacctgc	taccacttcg	900
	acctgcactg	ggagcaccac	cgctggccct	ttgccccctg	gtgggagctg	cccaactgcc	960

```

gccgcctgtc tggccgaggt ctggttcttg cctagctgga cacactgcag tgggccctgc 1020
tgccagctgg gcatgcaggt tgtggcagga ctgggtgagg tgaaaagctg caggcgctgc 1080
tgccggacac gctgcatggg ctaccctgtg tagctgccgc cactagggga gggggtttgt 1140
agctgtcgag cttgc

```

5

Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pGEM-Teasy (Promega) kloniert und der Klon pGKETO2 erhalten.

- 10 Sequenzierung des Klons pGKETO2 mit dem T7- und dem SP6-Primer bestätigte eine Sequenz, die sich lediglich in den drei Codons 73, 114 und 119 in je einer Base von der publizierten Sequenz X86782 unterscheidet. Diese Nukleotidaustausche wurden in einem unabhängigen Amplifikationsexperiment reproduziert und repräsentieren somit die Nukleotidsequenz im verwendeten *Haematococcus pluvialis* Stamm 192.80.

15

Dieser Klon wurde für die Klonierung in den unter Beispiel 1 beschriebenen Expressionsvektor verwendet. Die Klonierung erfolgte analog wie in Beispiel 1 beschrieben. Die Transformation der *E. coli* Stämme, deren Kultivierung und die Analyse des Carotinoidprofils erfolgte wie in Beispiel 3 beschrieben.

20

Abbildung 2 zeigt die chromatographische Analyse einer Probe erhalten aus einem mit diesem Expressionsvektor und pMCL-CrtYIBZ/idi/gps transformierten *E. coli*-Stamm. Unter Verwendung einer Ketolase aus *Haematococcus pluvialis*, wie beispielsweise in EP 725137 beschrieben, eluieren mit zunehmender Retentionszeit Astaxanthin (Peak 1), Adonixanthin (Peak 2) und nicht umgesetztes Zeaxanthin (Peak 3). Dieses Carotinoidprofil wurde bereits in EP 0725137 beschrieben.

25

Tabelle 1 zeigt einen Vergleich der bakteriell produzierten Carotinoidmengen:

30

Tabelle 1: Vergleich der bakteriellen Ketocarotinoid-Synthese bei Verwendung zweier verschiedener Ketolasen, der erfindungsgemäßen Ketolase aus *Nostoc* sp. Strain PCC7120 (Beispiel 3) und der Ketolase aus *Haematococcus pluvialis* als Vergleichsbeispiel (Beispiel 3.1). Carotinoidmengen sind in ng/ ml Kulturflüssigkeit angegeben.

Ketolase aus	Astaxanthin	Adonirubin	Adonixanthin	Canthaxanthin	Zeaxanthin
<i>Haematococcus pluvialis</i> <i>Flotow</i> em. Wille (Vergleichsbeispiel)	13		102		738
<i>Nostoc</i> sp. Strain PCC7120	491	186		120	

Die erfindungsgemäße Expression der Ketolase aus *Nostoc* sp. Strain PCC7120 führt zu einem Carotinoidmuster, welches sich von dem Carotinoidmuster nach Expression einer Ketolase aus *Haematococcus pluvialis* deutlich unterscheidet. Während die Ketolase aus dem Stand der Technik nur sehr unvollständig das gewünschte Ketocarotinoid Astaxanthin liefert, ist Astaxanthin bei der Verwendung der erfindungsgemäßen Ketolase das Hauptprodukt. Im erfindungsgemäßen Verfahren treten hydroxylierte Nebenprodukte in einer deutlich geringeren Menge auf.

Beispiel 4:

- Herstellung von Expressionsvektoren zur konstitutiven Expression der *Nostoc* sp. PCC 7120 Ketolase in *Lycopersicon esculentum* und *Tagetes erecta*.

Die Expression der Ketolase aus *Nostoc* in *L. esculentum* und in *Tagetes erecta* erfolgte unter Kontrolle des konstitutiven Promoters FNR (Ferredoxin NADPH Oxidoreductase) aus *Arabidopsis thaliana*. Die Expression erfolgte mit dem Transitpeptid rbcS aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715).

Das DNA Fragment, das die FNR Promotorregion -635 bis -1 aus *Arabidopsis thaliana* beinhaltet, wurde mittels PCR unter Verwendung genomischer DNA (nach Standardmethoden aus *Arabidopsis thaliana* isoliert) sowie der Primer FNR-1 (SEQ ID No.38) und FNR-2 (SEQ ID No. 39) hergestellt.

Die PCR-Bedingungen waren die folgenden:

- Die PCR zur Amplifikation der DNA, die das FNR-Promotorfragment FNR1-2 (-635 bis -1) beinhaltet, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

- 100 ng genomischer DNA aus *A.thaliana*
- 0.25 mM dNTPs
- 0.2 mM FNR-1 (SEQ ID No. 38)
- 0.2 mM FNR-2 (SEQ ID No. 39)
- 5 µl 10X PCR-Puffer (Stratagene)
- 0.25 µl Pfu Polymerase (Stratagene)
- 28.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X	94°C	2 Minuten
35X	94°C	1 Minute

50°C 1 Minute  
72°C 1 Minute  
1X 72°C 10 Minuten

- 5 Das 653 bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pFNR erhalten.

- 10 Sequenzierung des Klon pFNR bestätigte eine Sequenz, die mit einem Sequenzabschnitt auf Chromosom 5 von *Arabidopsis thaliana* (Datenbankeintrag AB011474) von Position 70127 bis 69493 übereinstimmt. Das Gen beginnt bei Basenpaar 69492 und ist mit "Ferredoxin-NADP+ Reductase" annotiert.

15 Dieser Klon heisst pFNR und wurde daher für die Klonierung in den Expressionsvektor pJIT117 (Guerineau et al. 1988, Nucl. Acids Res. 16: 11380) verwendet.

Die Klonierung erfolgte durch Isolierung des 635 bp SacI-HindIII Fragmentes aus pFNR und Ligierung in den SacI-HindIII geschnittenen Vektor pJIT117. Der Klon, der den Promoter FNR anstelle des ursprünglichen Promoters d35S enthält, heisst pJITFNR.

- 20 Zur Herstellung einer Expressionskassette pJFNRNOST wurde das 805 bp SphI-Fragment NOSTF-G (in Beispiel 1 beschrieben) in den SphI geschnittenen Vektor pJITFNR kloniert. Der Klon, der das Fragment NOSTF-G in der korrekten Orientierung als N-terminale Fusion mit dem rbcS Transitpeptid enthält, heisst pJFNRNOST.

- 25 Die Herstellung einer Expressionskassette für die Agrobacterium vermittelte Transformation der Ketolase aus *Nostoc* in *L. esculentum* erfolgte unter der Verwendung des binären Vektors pSUN3 (WO02/00900).

- 30 Zur Herstellung des Expressionsvektors pS3FNRNOST wurde das 2.4 Kb SacI-XhoI Fragment (partielle SacI Hydrolyse) aus pJFNRNOST mit dem SacI-XhoI geschnittenen Vektor pSUN3 ligiert (Abbildung 3, Konstruktkarte). In der Abbildung 3 beinhaltet Fragment *FNR-Promotor* den FNR Promotor (655 bp), Fragment *rbcS TransitPeptid* das rbcS Transitpeptid aus Erbse (204 bp), Fragment *Nost Ketolase* (799 bp) die gesamte Primärsequenz, kodierend für die *Nostoc* Ketolase, Fragment *35S Terminator* (761 bp) das Polyadenylierungssignal von CaMV.

- 35 Die Herstellung einer Expressionskassette für die *Agrobacterium*-vermittelte Transformation des Expressionsvektor mit der Ketolase aus *Nostoc* in *Tagetes erecta* erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).

Zur Herstellung des Tagetes-Expressionsvektors pS5FNRNOST wurde das 2.4 Kb SacI-XhoI Fragment (partielle SacI Hydrolyse) aus pJFNRNOST mit dem SacI-XhoI geschnittenen Vektor pSUN5 ligiert (Abbildung 4, Konstruktkarte). In der Abbildung 4 beinhaltet Fragment *FNR Promotor* den duplizierten FNR Promotor (655 bp), Fragment *rbcS Transit Peptid* das *rbcS* Transitpeptid aus Erbse (204 bp), Fragment *Nost Ketolase* (799 bp) die gesamte Primärsequenz, kodierend für die *Nostoc* Ketolase, Fragment *35S Terminator* (761 bp) das Polyadenylierungssignal von CaMV.

Beispiel 5:

Herstellung von Expressionsvektoren zur blütenspezifischen Expression der *Nostoc* sp. *PCC 7120* Ketolase in *Lycopersicon esculentum* und *Tagetes erecta*.

Die Expression der Ketolase aus *Nostoc* in *L. esculentum* und *Tagetes erecta* erfolgte mit dem Transitpeptid *rbcS* aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715). Die Expression erfolgte unter Kontrolle einer modifizierten Version AP3P des blütenspezifischen Promoters AP3 aus *Arabidopsis thaliana* (AL132971: Nukleotidregion 9298-10200; Hill et al. (1998) Development 125: 1711-1721).

Das DNA Fragment, das die AP3 Promoterregion -902 bis +15 aus *Arabidopsis thaliana* beinhaltet, wurde mittels PCR unter Verwendung genomischer DNA (nach Standardmethoden aus *Arabidopsis thaliana* isoliert) sowie der Primer AP3-1 (SEQ ID No.41) und AP3-2 (SEQ ID No. 42) hergestellt.

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der DNA, die das AP3-Promoterfragment (-902 bis +15) beinhaltet, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

- 100 ng genomischer DNA aus *A.thaliana*
- 0.25 mM dNTPs
- 0.2 mM AP3-1 (SEQ ID No. 41)
- 0.2 mM AP3-2 (SEQ ID No. 42)
- 5 µl 10X PCR-Puffer (Stratagene)
- 0.25 µl Pfu Polymerase (Stratagene)
- 28.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X    94°C    2 Minuten

35X 94°C 1 Minute  
50°C 1 Minute  
72°C 1 Minute  
1X 72°C 10 Minuten

5

Das 929 Bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pAP3 erhalten.

10

Sequenzierung des Klon pAP3 bestätigte eine Sequenz, die sich lediglich in durch eine Insertion (ein G in Position 9765 der Sequenz AL132971) und einen Basenaustausch (ein G statt ein A in Position 9726 der Sequenz AL132971) von der publizierten AP3 Sequenz (AL132971, Nukleotidregion 9298-10200) unterscheidet. Diese Nukleotidunterschiede wurden in einem unabhängigen Amplifikationsexperiment reproduziert und repräsentieren somit die tatsächliche Nukleotidsequenz in den verwendeten *Arabidopsis thaliana* Pflanzen.

15

Die modifizierte Version AP3P wurde mittels rekombinanter PCR unter Verwendung des Plasmids pAP3 hergestellt. Die Region 10200 - 9771 wurde mit den Primern AP3-1 (SEQ ID No. 41) und Primern AP3-4 (SEQ ID No. 44) amplifiziert (Amplifikat A1/4), die Region 9526-9285 wurde mit den AP3-3 (SEQ ID No. 43) und AP3-2 (SEQ ID No. 42) amplifiziert (Amplifikat A2/3).

20

Die PCR-Bedingungen waren die folgenden:

25

Die PCR-Reaktionen zur Amplifikation der DNA-Fragmente, die die Regionen Region 10200 - 9771 und Region 9526-9285 des AP3 Promoters beinhalten, erfolgte in 50 µl Reaktionsansätzen, in denen enthalten war:

30

- 100 ng AP3 Amplifikat (oben beschrieben)
- 0.25 mM dNTPs
- 0.2 mM sense Primer (AP3-1 SEQ ID No. 41 bzw. AP3-3 SEQ ID No. 43)
- 0.2 mM antisense Primer (AP3-4 SEQ ID No. 44 bzw. AP3-2 SEQ ID No. 42)
- 5 µl 10X PCR-Puffer (Stratagene)
- 0.25 µl Pfu Taq Polymerase (Stratagene)
- 28.8 µl Aq. Dest.

35

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X 94°C 2 Minuten  
35X 94°C 1 Minute

50°C 1 Minute  
72°C 1 Minute  
1X 72°C 10 Minuten

5 Die rekombinante PCR beinhaltet Annealing der sich über eine Sequenz von 25 Nukleotiden überlappenden Amplifikate A1/4 und A2/3, Vervollständigung zu einem Doppelstrang und anschließende Amplifizierung. Dadurch entsteht eine modifizierte Version des AP3 Promoters, AP3P, in dem die Positionen 9670 - 9526 deletiert sind. Die Denaturierung (5 min bei 95°C) und Annealing (langsame Abkühlung bei Raumtemperatur auf 40°C) beider Amplifikate A1/4 und  
10 A2/3 erfolgte in einem 17.6 µl Reaktionsansatz, in dem enthalten war:

- 0.5 µg A1/4 Amplifikat
- 0.25 µg A2/3 Amplifikat

15 Das Auffüllen der 3'-Enden (30 min bei 30°C) erfolgte in einem 20 µl Reaktionsansatz, in dem enthalten war:

- 17.6 µl A1/4 und A2/3-Annealingsreaktion (hergestellt wie oben beschrieben)
- 50 µM dNTPs
- 20 - 2 µl 1X Klenow Puffer
- 2U Klenow Enzym

Die Nukleinsäure kodierend für die modifizierte Promoterversion AP3P wurde mittels PCR unter Verwendung eines sense spezifischen Primers (AP3-1 SEQ ID No. 41) und eines antisense  
25 spezifischen Primers (AP3-2 SEQ ID No. 42) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

30 Die PCR zur Amplifikation des AP3P Fragmentes erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

- 1 µl Annealingsreaktion (hergestellt wie oben beschrieben)
- 0.25 mM dNTPs
- 0.2 mM AP3-1 (SEQ ID No. 41)
- 35 - 0.2 mM AP3-2 (SEQ ID No. 42)
- 5 µl 10X PCR-Puffer (Stratagene)
- 0.25 µl Pfu Taq Polymerase (Stratagene)



- 28.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

5	1X	94°C	2 Minuten
	35X	94°C	1 Minute
		50°C	1 Minute
		72°C	1 Minute
	1X	72°C	10 Minuten

10

Die PCR-Amplifikation mit SEQ ID No. 41 (AP3-1) und SEQ ID No. 42 (AP3-2) resultierte in einem 783 Bp Fragment, das für die modifizierte Promoterversion AP3P kodiert. Das Amplifikat wurde in den Klonierungsvektor pCR2.1 (Invitrogen) kloniert und das Plasmid pAP3P erhalten. Sequenzierungen mit den Primern T7 und M13 bestätigten eine zur Sequenz AL132971, Region 10200-9298 identische Sequenz, wobei die interne Region 9285 - 9526 deletiert wurde. Dieser Klon wurde daher für die Klonierung in den Expressionsvektor pJIT117 (Guerineau et al. 1988, Nucl. Acids Res. 16: 11380) verwendet.

15

Die Klonierung erfolgte durch Isolierung des 783 Bp SacI-HindIII Fragmentes aus pAP3P und Ligierung in den SacI-HindIII geschnittenen Vektor pJIT117. Der Klon, der den Promoter AP3P anstelle des ursprünglichen Promoters d35S enthält, heisst pJITAP3P. Zur Herstellung einer Expressionskassette pJAP3NOST wurde das 805 Bp SphI-Fragment NOSTF-G (in Beispiel 1 beschrieben) in den SphI geschnittenen Vektor pJITAP3P kloniert. Der Klon, der das Fragment NOSTF-G in der korrekten Orientierung als N-terminale Fusion mit dem rbcS Transitpeptid enthält, heisst pJAP3PNOST.

25

Die Herstellung eines Expressionsvektors für die Agrobacterium-vermittelte Transformation der AP3P-kontrollierten Ketolase aus Nostoc in *L. esculentum* erfolgte unter der Verwendung des binären Vektors pSUN3 (WO02/00900).

30

Zur Herstellung des Expressionsvektors pS3AP3PNOST wurde das 2.6 KB bp SacI-XhoI Fragment (partielle SacI Hydrolyse) aus pJAP3NOST mit dem SacI-XhoI geschnittenen Vektor pSUN3 ligiert (Abbildung 5, Konstruktkarte). In der Abbildung 5 beinhaltet Fragment AP3P den modifizierten AP3P Promoter (783 bp), Fragment rbcS das rbcS Transitpeptid aus Erbse (207 bp), Fragment NOSTF-G (792 bp) die gesamte Primärsequenz kodierend für die Nostoc Ketolase, Fragment term (795 bp) das Polyadenylierungssignal von CaMV.

35

Die Herstellung einer Expressionsvektors für die Agrobacterium-vermittelte Transformation der AP3P-kontrollierten Ketolase aus Nostoc in Tagetes erecta erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).

- 5 Zur Herstellung des Expressionsvektors pS5AP3PNOST wurde das 2.6 KB bp SacI-XhoI (partielle SacI Hydrolyse) Fragment aus pS5AP3PNOST mit dem SacI-XhoI geschnittenen Vektor pSUN5 ligiert (Abbildung 6, Konstruktkarte). In der Abbildung 6 beinhaltet Fragment AP3P den modifizierten AP3P Promoter (783 bp), Fragment rbcS das rbcS Transitpeptid aus Erbse (207 bp), Fragment NOSTF-G (792 bp) die gesamte Primärsequenz codierend für die Nostoc Ketolase, Fragment term (795 bp) das Polyadenylierungssignal von CaMV.
- 10

Beispiel 6:

Herstellung transgener Lycopersicon esculentum Pflanzen

- 15 Transformation und Regeneration von Tomatenpflanzen erfolgte nach der publizierten Methode von Ling und Mitarbeitern (Plant Cell Reports (1998), 17:843-847). Für die Varietät Microtom wurde mit höherer Kanamycin-Konzentration (100mg/L) selektioniert.

- 20 Als Ausgangsexplantat für die Transformation dienten Kotyledonen und Hypokotyle sieben bis zehn Tage alter Keimlinge der Linie Microtom. Für die Keimung wurde das Kulturmedium nach Murashige und Skoog (1962: Murashige and Skoog, 1962, Physiol. Plant 15, 473-) mit 2% Saccharose, pH 6,1 verwendet. Die Keimung fand bei 21°C bei wenig Licht (20 - 100 µE) statt. Nach sieben bis zehn Tagen wurden die Kotyledonen quer geteilt und die Hypokotyle in ca. 5 - 10 mm lange Abschnitte geschnitten und auf das Medium MSBN (MS, pH 6,1, 3% Saccharose + 25 1 mg/l BAP, 0,1 mg/l NAA) gelegt, das am Vortag mit suspensionskultivierten Tomatenzellen beschickt wurde. Die Tomatenzellen wurden luftblasenfrei mit sterilem Filterpapier abgedeckt. Die Vorkultur der Explantate auf dem beschriebenen Medium erfolgte für drei bis fünf Tage. Zellen des Stammes Agrobacterium tumefaciens LBA4404 wurden einzeln mit den Plasmiden pS3FNRNOST und pS3AP3NOST transformiert. Von den einzelnen mit den Binärvektoren 30 pS3FNRNOST und pS3AP3NOST transformierten Agrobacterium-Stämmen wurde jeweils eine Übernachtkultur in YEB Medium mit Kanamycin (20 mg/l) bei 28 °C kultiviert und die Zellen zentrifugiert. Das Bakterienpellet wurde mit flüssigem MS Medium (3% Saccharose, pH 6,1) resuspendiert und auf eine optische Dichte von 0,3 (bei 600 nm) eingestellt. Die vorkultivierten Explantate wurden in die Suspension überführt und für 30 Minuten bei Zimmertemperatur unter 35 leichtem Schütteln inkubiert. Anschließend wurden die Explantate mit sterilem Filterpapier getrocknet und für die dreitägige Co-Kultur (21°C) auf ihr Vorkulturmedium zurück gelegt.

Nach der Co-kultur wurden die Explantate auf MSZ2 Medium (MS pH 6,1 + 3% Saccharose, 2 mg/l Zeatin, 100 mg/l Kanamycin, 160 mg/l Timentin) transferiert und für die selektive Regenera-

tion bei 21°C unter Schwachlicht Bedingungen (20 - 100  $\mu$ E, Lichtrhythmus 16h/8h) aufbewahrt. Aller zwei bis drei Wochen erfolgte der Transfer der Explantate bis sich Sprosse bilden. Kleine Sprosse konnten vom Explantat abgetrennt werden und auf MS (pH 6,1 + 3% Saccharose) 160 mg/l Timentin, 30 mg/l Kanamycin, 0,1 mg/l IAA bewurzelt werden. Bewurzelte Pflanzen wurden  
5 ins Gewächshaus überführt.

Gemäß der oben beschriebenen Transformationsmethode wurden mit folgenden Expressionskonstrukten folgende Linien erhalten:

10 Mit pS3FNRNOST wurde erhalten: ms 101-1, ms101-2, ms101-3

Mit pS3AP3NOST wurde erhalten: ms 102-1, ms102-2, ms102-3

Beispiel 7:

15 Herstellung transgener Tagetes Pflanzen

Tagetessamen werden sterilisiert und auf Keimungsmedium (MS-Medium; Murashige and Skoog, Physiol. Plant. 15(1962), 473-497) pH 5,8, 2% Saccharose) aufgelegt. Die Keimung erfolgt in einem Temperatur/Licht/Zeitintervall von 18-28°C/20-200  $\mu$ E/3 - 16 Wochen, bevorzugt  
20 jedoch bei 21°C, 20-70  $\mu$ E, für 4-8 Wochen.

Alle Blätter der sich bis dahin entwickelten in vitro Pflanzen werden geerntet und quer zur Mittelrippe geschnitten. Die dadurch entstehenden Blattexplantate mit einer Größe von 10 - 60 mm<sup>2</sup> werden im Verlaufe der Präparation in flüssigem MS - Medium bei Raumtemperatur für maximal  
25 2 h aufbewahrt.

Ein beliebiger Agrobakterium tumefaciens Stamm, bevorzugt aber ein supervirulenter Stamm, wie z.B. EHA105 mit einem entsprechenden Binärplasmid, das ein Selektionsmarkergen (bevorzugt *bar* oder *pat*) sowie ein oder mehrere Trait- oder Reportergene tragen kann wird  
30 (pS5FNRNOST und pS5AP3NOST), über Nacht angezogen und für die Co-Kultivierung mit dem Blattmaterial verwendet. Die Anzucht des Bakterienstammes kann wie folgt erfolgen: Eine Einzelkolonie des entsprechenden Stammes wird in YEB (0,1 % Hefeextrakt, 0,5 % Rindfleischextrakt, 0,5 % Pepton, 0,5 % Saccharose, 0,5 % Magnesiumsulfat x 7 H<sub>2</sub>O) mit 25 mg/l Kanamycin angeimpft und bei 28°C für 16 bis 20 h angezogen. Anschließend wird die Bakteriensuspension  
35 durch Zentrifugation bei 6000 g für 10 min geerntet und derart in flüssigem MS Medium resuspendiert, daß eine OD<sub>600</sub> von ca. 0,1 bis 0,8 entstand. Diese Suspension wird fuer die C-Kultivierung mit dem Blattmaterial verwendet.

Unmittelbar vor der Co-Kultivierung wird das MS-Medium, in dem die Blätter aufbewahrt worden sind, durch die Bakteriensuspension ersetzt. Die Inkubation der Blättchen in der Agrobakterien-suspension erfolgte für 30 min unter leichtem Schütteln bei Raumtemperatur. Anschließend werden die infizierten Explantate auf ein mit Agar (z.B. 0,8 % Plant Agar (Duchefa, NL) verfestig-  
5 tes MS-Medium mit Wachstumsregulatoren, wie beispielsweise 3 mg/l Benzylaminopurin (BAP) sowie 1 mg/l Indolylessigsäure (IAA) aufgelegt. Die Orientierung der Blätter auf dem Medium ist bedeutungslos. Die Kultivierung der Explantate findet für 1 bis 8 Tage, bevorzugt aber für 6 Ta-  
ge statt, dabei können folgende Bedingungen angewendet werden: Lichtintensität: 30 – 80  $\mu\text{Mol/m}^2 \times \text{sec}$ , Temperatur: 22 – 24°C, hell/dunkel Wechsel von 16/8 Stunden. Anschließend  
10 werden die co-kultivierten Explantate auf frisches MS-Medium, bevorzugt mit den gleichen Wachstumsregulatoren übertragen, wobei dieses zweite Medium zusätzlich ein Antibiotikum zur Unterdrückung des Bakterienwachstums enthält. Timentin in einer Konzentration von 200 bis 500 mg/l ist für diesen Zweck sehr geeignet. Als zweite selektive Komponente wird eine für die  
15 Selektion des Transformationserfolges eingesetzt. Phosphinothricin in einer Konzentration von 1 bis 5 mg/l selektiert sehr effizient, aber auch andere selektive Komponenten gemäß des zu ver-wendenden Verfahrens sind denkbar.

Nach jeweils ein bis drei Wochen erfolgt der Transfer der Explantate auf frisches Medium bis  
20 sich Sproßknospen und kleine Sprosse entwickeln, die dann auf das gleiche Basalmedium ein-schließlich Timentin und PPT oder alternative Komponenten mit Wachstumsregulatoren, näm-lich z.B. 0,5 mg/l Indolylbuttersäure (IBA) und 0,5 mg/l Gibberillinsäure  $\text{GA}_3$ , zur Bewurzelung  
übertragen werden. Bewurzelte Sprosse können ins Gewächshaus überführt werden.

Zusätzlich zu der beschriebenen Methode sind folgende vorteilhafte Modifikationen möglich:

25 Bevor die Explantate mit den Bakterien infiziert werden, können sie für 1 bis 12 Tage, bevorzugt 3 - 4, auf das oben beschriebene Medium für die Co-Kultur vorinkubiert werden. Anschließend erfolgt die Infektion, Co-Kultur und selektive Regeneration wie oben beschrieben.

30 Der pH Wert für die Regeneration (normalerweise 5,8) kann auf pH 5,2 gesenkt werden. Da-durch wird die Kontrolle des Agrobakterienwachstums verbessert.

Die Zugabe von  $\text{AgNO}_3$  (3 - 10 mg/l) zum Regenerationsmedium verbessert den Zustand der Kultur einschließlich der Regeneration selbst.

35 Komponenten, die die Phenolbildung reduzieren und dem Fachmann bekannt sind, wie z.B. Zitronensäure, Ascorbinsäure, PVP u.v.a.m., wirken sich positiv auf die Kultur aus.

Für das gesamte Verfahren kann auch flüssiges Kulturmedium Verwendung finden. Die Kultur kann auch auf handelsüblichen Trägern, die auf dem flüssigen Medium positioniert werden inkubiert werden.

- 5 Gemäß der oben beschriebenen Transformationsmethode wurden mit folgenden Expressionskonstrukten folgende Linien erhalten:

Mit pS5FNRNOST wurde beispielsweise erhalten: ms 103-1, ms103-2, ms103-3, mit  
pS5AP3NOST wurde beispielsweise erhalten: ms 104-1, ms104-2, ms104-3

10

#### Beispiel 9

Charakterisierung der transgenen Pflanzenblüten

#### Beispiel 9.1

15

Trennung von Carotinoidestern in Blütenblättern transgener Pflanzen

#### Allgemeine Arbeitsvorschrift:

20

Die Blütenblätter der transgenen Pflanzen werden in flüssigem Stickstoff gemörsert und das Petalenpulver (etwa 40 mg) mit 100% Aceton extrahiert (dreimal je 500 µl). Das Lösungsmittel wird evaporiert und die Carotinoide in 100-200 µl Petrolether/Aceton (5:1, v/v) resuspendiert.

25

Die Carotinoide werden in konzentrierter Form mittels Dünnschicht-Chromatographie (TLC) auf Silica60 F254- Platten (Merck) in einem organischen Laufmittel (Petrolether/Aceton; 5:1) entsprechend ihrer Phobizität aufgetrennt. Gelbe (Xanthophyllester), rote (Ketocarotinoidester) und orange Banden (Mischung aus Xanthophyll- und Ketocarotinoidestern) auf der TLC werden ausgekratzt.

30

Die an Silica gebundenen Carotinoide werden dreimal mit 500 µl Aceton eluiert, das Lösungsmittel evaporiert und die Carotinoide mittels HPLC aufgetrennt und identifiziert.

35

Mittels einer C30-reverse phase-Säule kann zwischen Mono- und Diestern der Carotinoide unterschieden werden. HPLC-Laufbedingungen waren nahezu identisch mit einer publizierten Methode (Frazer et al.(2000), Plant Journal 24(4): 551-558). Folgende Verfahrensbedingungen wurden eingestellt.

Trennsäule: ProntoSIL C30-Säule, 250 x 4,6 mm, (Bischoff, Leonberg)

Flussrate: 1.0 ml/min

Eluenten: Laufmittel A - 100% Methanol

Laufmittel B - 80% Methanol, 0.2% Ammoniumacetat  
Laufmittel C - 100% t-Butyl-methylether

Gradientprofil:

Zeit	Flussrate	% Laufmittel A	% Laufmittel B	% Laufmittel C
12.0	1.0	95.0	5.0	0
12.1	1.0	80.0	5.0	15.0
22.0	1.0	76.0	5.0	19.0
22.0	1.0	66.5	5.0	28.5
38.0	1.0	15.0	5.0	80.0
45.0	1.0	95.0	5.0	0
46.0	1.0	95.0	5.0	0
46.1	1.0	95.0	5.0	0

Detektion: 300 - 500 nm

Eine Identifizierung der Carotinoide ist aufgrund der UV-VIS-Spektren möglich.

Petalenmaterial der transgenen Tomatenpflanzen wird gemörsert und mit Aceton extrahiert. Extrahierte Carotinoide werden mittels TLC aufgetrennt. In den Linien können Mono- und Diester von Ketocarotinoiden detektiert werden; die Monoester sind in deutlich geringerer Konzentration als die Diester vorhanden.

#### Beispiel 10

Enzymatische Hydrolyse von Carotinoidestern und Identifizierung der Carotinoide

#### Allgemeine Arbeitsvorschrift

Gemörsertes Petalenmaterial (30-100 mg Frischgewicht) wird mit 100% Aceton (dreimal 500ul; jeweils etwa 15 Minuten schütteln) extrahiert. Das Lösungsmittel wird evaporiert. Carotinoide werden anschließend in 495 ul Aceton aufgenommen, 4,95 ml Kalium-phosphatpuffer (100 mM, pH7.4) zugegeben und gut gemischt. Danach erfolgt die Zugabe von ca. 17 mg Bile-Salze (Sigma) und 149 µl einer NaCl/CaCl<sub>2</sub>-Lösung (3M NaCl und 75 mM CaCl<sub>2</sub>). Die Suspension wird für 30 Minuten bei 37C inkubiert. Für die enzymatische Hydrolyse der Carotinoidester wird 595 µl einer Lipaselösung (50 mg/ml Lipase Typ7 von Candida rugosa(Sigma)) zugegeben und unter Schütteln bei 37C inkubiert. Nach etwa 21 Stunden erfolgte nochmals eine Zugabe von 595 µl Lipase mit erneuter Inkubation von mindestens 5 Stunden bei 37C. Anschließend werden etwa

- ca. 700 mg  $\text{Na}_2\text{SO}_4 \cdot 10\text{H}_2\text{O}$  in der Lösung gelöst. Nach Zugabe von 1800  $\mu\text{l}$  Petrolether werden die Carotinoide durch kräftig Mischen in die organische Phase extrahiert. Dieses Ausschütteln wird solange wiederholt, bis die organische Phase farblos bleibt. Die Petroletherfraktionen werden vereinigt und der Petrolether evaporiert. Freie Carotinoide werden in 100-120  $\mu\text{l}$  Aceton aufgenommen. Mittels HPLC und C30-reverse phase-Säule können freie Carotinoide aufgrund von Retentionszeit und UV-VIS-Spektren identifiziert werden.
- 5

Verfahren zur Herstellung von Ketocarotinoiden in genetisch veränderten Organismen

Zusammenfassung

- 5 Die vorliegende Erfindung betrifft ein Verfahren zur Herstellung von Ketocarotinoiden durch Kultivierung von genetisch veränderten Organismen, die im Vergleich zum Wildtyp eine veränderte Ketolase-Aktivität aufweisen, die genetisch veränderten Organismen, sowie deren Verwendung als Nahrungs- und Futtermittel und zur Herstellung von Ketocarotinoidextrakten.



Nostoc.ST25.txt  
SEQUENCE LISTING

<110> SunGene GmbH & Co. KGaA

<120> Verfahren zur Herstellung von Ketocarotinoiden in genetisch veränderten Organismen

<130> 20020636

<160> 51

<170> PatentIn version 3.1

<210> 1

<211> 777

<212> DNA

<213> Nostoc sp. Strain PCC7120

<220>

<221> CDS

<222> (1)..(777)

<223>

ctg gtt cag tgt caa cca tca tct ctg cat tca gaa aaa ctg gtg tta	48
Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu	
1 5 10 15	
ttg tca tcg aca atc aga gat gat aaa aat att aat aag ggt ata ttt	96
Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe	
20 25 30	
att gcc tgc ttt atc tta ttt tta tgg gca att agt tta atc tta tta	144
Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu	
35 40 45	
ctc tca ata gat aca tcc ata att cat aag agc tta tta ggt ata gcc	192
Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala	
50 55 60	
atg ctt tgg cag acc ttc tta tat aca ggt tta ttt att act gct cat	240
Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His	
65 70 75 80	
gat gcc atg cac ggc gta gtt tat ccc aaa aat ccc aga ata aat aat	288
Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn	

Nostoc.ST25.txt

85

90

95

ttt	ata	ggt	aag	ctc	act	cta	atc	ttg	tat	gga	cta	ctc	cct	tat	aaa	336
Phe	Ile	Gly	Lys	Leu	Thr	Leu	Ile	Leu	Tyr	Gly	Leu	Leu	Pro	Tyr	Lys	
			100					105					110			
gat	tta	ttg	aaa	aaa	cat	tgg	tta	cac	cac	gga	cat	cct	ggt	act	gat	384
Asp	Leu	Leu	Lys	Lys	His	Trp	Leu	His	His	Gly	His	Pro	Gly	Thr	Asp	
		115					120					125				
tta	gac	cct	gat	tat	tac	aat	ggt	cat	ccc	caa	aac	ttc	ttt	ctt	tgg	432
Leu	Asp	Pro	Asp	Tyr	Tyr	Asn	Gly	His	Pro	Gln	Asn	Phe	Phe	Leu	Trp	
	130					135					140					
tat	cta	cat	ttt	atg	aag	tct	tat	tgg	cga	tgg	acg	caa	att	ttc	gga	480
Tyr	Leu	His	Phe	Met	Lys	Ser	Tyr	Trp	Arg	Trp	Thr	Gln	Ile	Phe	Gly	
	145				150					155					160	
tta	gtg	atg	att	ttt	cat	gga	ctt	aaa	aat	ctg	gtg	cat	ata	cca	gaa	528
Leu	Val	Met	Ile	Phe	His	Gly	Leu	Lys	Asn	Leu	Val	His	Ile	Pro	Glu	
				165					170					175		
aat	tta	att	ata	ttt	tgg	atg	ata	cct	tct	att	tta	agt	tca	gta		576
Asn	Leu	Ile	Ile	Phe	Trp	Met	Ile	Pro	Ser	Ile	Leu	Ser	Ser	Val		
		180					185					190				
caa	cta	ttt	tat	ttt	ggt	aca	ttt	ttg	cct	cat	aaa	aag	cta	gaa	ggt	624
Gln	Leu	Phe	Tyr	Phe	Gly	Thr	Phe	Leu	Pro	His	Lys	Lys	Leu	Glu	Gly	
		195					200					205				
ggt	tat	act	aac	ccc	cat	tgt	gcg	cg	agt	atc	cca	tta	cct	ctt	ttt	672
Gly	Tyr	Thr	Asn	Pro	His	Cys	Ala	Arg	Ser	Ile	Pro	Leu	Pro	Leu	Phe	
	210					215					220					
tgg	tct	ttt	gtt	act	tgt	tat	cac	ttc	ggc	tac	cac	aag	gaa	cat	cac	720
Trp	Ser	Phe	Val	Thr	Cys	Tyr	His	Phe	Gly	Tyr	His	Lys	Glu	His	His	
	225				230				235						240	
gaa	tac	cct	caa	ctt	cct	tgg	tgg	aaa	tta	cct	gaa	gct	cac	aaa	ata	768
Glu	Tyr	Pro	Gln	Leu	Pro	Trp	Trp	Lys	Leu	Pro	Glu	Ala	His	Lys	Ile	
				245					250					255		
tct	tta	taa														777
Ser	Leu															

<210> 2

<211> 258

<212> PRT

<213> Nostoc sp. Strain PCC7120

<400> 2

Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu  
1 5 10 15

Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe  
20 25 30

Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu  
Seite 2

Nostoc.ST25.txt

35

40

45

Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala  
50 55 60

Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His  
65 70 75 80

Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn  
85 90 95

Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys  
100 105 110

Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp  
115 120 125

Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp  
130 135 140

Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly  
145 150 155 160

Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu  
165 170 175

Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val  
180 185 190

Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly  
195 200 205

Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe  
210 215 220

Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His  
225 230 235 240

Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile  
245 250 255

Ser Leu

<210> 3

<211> 789

<212> DNA

<213> Nostoc punctiforme

Nostoc.ST25.txt

**<220>**

**<221> CDS**

<222> (1)..(789)

**<223>**

[illegible]

Nostoc.ST25.txt

gct tgc tac cac ttt ggt tat cat gaa gaa cat cat gag tat ccc cat 720  
 Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
 225 230 235 240

gta cct tgg tgg caa ctt cca tct gta tat aag cag aga gta ttc aac 768  
 Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn  
 245 250 255

aat tca gta acc aat tcg taa 789  
 Asn Ser Val Thr Asn Ser  
 260

<210> 4  
 <211> 262  
 <212> PRT  
 <213> Nostoc punctiforme

Met Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln  
 1 5 10 15

Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val  
 20 25 30

Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn  
 35 40 45

Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln  
 50 55 60

Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His  
 65 70 75 80

Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser  
 85 90 95

Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys  
 100 105 110

Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp  
 115 120 125

Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe  
 130 135 140

Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu  
 145 150 155 160

Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile  
 165 170 175

Nostoc.ST25.txt

Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr  
180 185 190

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr  
195 200 205

Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile  
210 215 220

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
225 230 235 240

Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn  
245 250 255

Asn Ser Val Thr Asn Ser  
260

<210> 5

<211> 762

<212> DNA

<213> Nostoc punctiforme

<220>

<221> CDS

<222> (1)..(762)

<223>

Met Ile Gln Leu Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val	48
gta ctg aga agt aaa tct cag ttt aag ggg ctt ttc att gct att gtc	96
Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val	
att gtt agc gca tgg gtc att agc ctg agt tta tta ctt tcc ctt gac	144
Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu 45	
atc tca aag cta aaa ttt tgg atg tta ttg cct gtt ata cta tgg caa	192
Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln	
aca ttt tta tat acg gga tta ttt att aca tct cat gat gcc atg cat	240
Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His	
ggc gta gta ttt ccc caa aac acc aag att aat cat ttg att gga aca	288
Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr	

Nostoc.ST25.txt

85

90

95

ttg acc cta tcc ctt tat ggt ctt tta cca tat caa aaa cta ttg aaa	336
Leu Thr Leu Ser 100 Leu Tyr Gly Leu 105 Leu Pro Tyr Gln Lys 110 Leu Leu Lys	
aaa cat tgg tta cac cac cac aat cca gca agc tca ata gac ccg gat	384
Lys His Trp 115 Leu His His His Asn 120 Pro Ala Ser Ser 125 Ile Asp Pro Asp	
ttt cac aat ggt aaa cac caa agt ttc ttt gct tgg tat ttt cat ttt	432
Phe His Asn 130 Gly Lys His Gln 135 Ser Phe Phe Ala Trp 140 Tyr Phe His Phe	
atg aaa ggt tac tgg agt tgg ggg caa ata att gcg ttg act att att	480
Met Lys Gly Tyr Trp 150 Ser Trp Gly Gln Ile 155 Ala Leu Thr Ile Ile 160	
tat aac ttt gct aaa tac ata ctc cat atc cca agt gat aat cta act	528
Tyr Asn Phe Ala 165 Lys Tyr Ile Leu His 170 Ile Pro Ser Asp Asn 175 Leu Thr	
ttt tgg gtg cta ccc tcg ctt tta agt tca tta caa tta ttc tat	576
Phe Trp Val 180 Leu Pro Ser Leu 185 Ser Ser Leu Gln 190 Leu Phe Tyr	
ttt ggt act ttt tta ccc cat agt gaa cca ata ggg ggt tat gtt cag	624
Phe Gly Thr 195 Phe Leu Pro His Ser 200 Glu Pro Ile Gly 205 Tyr Val Gln	
cct cat tgt gcc caa aca att agc cgt cct att tgg tgg tca ttt atc	672
Pro His Cys Ala Gln Thr 215 Ile Ser Arg Pro Ile 220 Trp Trp Ser Phe Ile	
acg tgc tat cat ttt ggc tac cac gag gaa cat cac gaa tat cct cat	720
Thr Cys Tyr His Phe 230 Gly Tyr His Glu Glu 235 His His Glu Tyr Pro His 240	
att tct tgg tgg cag tta cca gaa att tac aaa gca aaa tag	762
Ile Ser Trp Trp 245 Gln Leu Pro Glu Ile 250 Tyr Lys Ala Lys	

<210> 6

<211> 253

<212> PRT

<213> Nostoc punctiforme

<400> 6

Met Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro  
1 5 10 15

Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val  
20 25 30

Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp  
35 40 45

Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln  
Seite 7

Nostoc.ST25.txt  
60

50

55

Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His  
65 70 75 80

Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr  
85 90 95

Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys  
100 105 110

Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp  
115 120 125

Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe  
130 135 140

Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile  
145 150 155 160

Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr  
165 170 175

Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr  
180 185 190

Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln  
195 200 205

Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile  
210 215 220

Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
225 230 235 240

Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys  
245 250

<210> 7

<211> 789

<212> DNA

<213> Künstliche Sequenz

<220>

<221> CDS

<222> (1)..(789)

<223>



**Nostoc.ST25.txt**

[illegible]

aat tca gta acc aat tcg taa  
 Asn Ser Val Thr Asn Ser  
 260

<210> 8

<211> 262

<212> PRT

<213> Künstliche Sequenz

<400> 8

Met Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln  
 1 5 10 15

Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val  
 20 25 30

Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn  
 35 40 45

Tyr Ala Lys Ile His Lys Trp Leu Ile Pro Ile Ala Ile Val Trp Gln  
 50 55 60

Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His  
 65 70 75 80

Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser  
 85 90 95

Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys  
 100 105 110

His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp  
 115 120 125

Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe  
 130 135 140

Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu  
 145 150 155 160

Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile  
 165 170 175

Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr  
 180 185 190

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr  
 195 200 205

Nostoc.ST25.txt

Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile  
210 215 220

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
225 230 235 240

Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn  
245 250 255

Asn Ser Val Thr Asn Ser  
260

<210> 9

<211> 789

<212> DNA

<213> Künstliche Sequenz

<220>

<221> CDS

<222> (1)..(789)

<223>

<400> 9 48

atg aat ttt tgt gat aaa cca gtt agc tat tat gtt gca ata gag caa  
Met Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln  
1 5 10 15

tta agt gct aaa gaa gat act gtt tgg ggg ctg gtg att gtc ata gta  
Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val  
20 25 30 96

ctt att agt ctt tgg gta gct agt ttg gct ttt tta cta gct att aat  
Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn  
35 40 45 144

tat gcc aaa gtc cca att tgg ttg ata cct att gca ata gtt tgg caa  
Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln  
50 55 60 192

atg ttc ctt tat aca ggg cta ttt att act gca cat gat gct atg cat  
Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His  
65 70 75 80 240

ggg tca gtt tat cgt aaa aat ccc aaa att aat aat ttt atc ggt tca  
Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser  
85 90 95 288

cta gct gta gcg ctt tac gct gtg ttt cca tat caa cag atg tta aag  
Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys  
100 105 110 336

aat cat tgc tta cat cat cgt cat cct gct agc gat tta gac cca gat  
Asn His Cys Leu His His Arg His Pro Ala Ser Asp Leu Asp Pro Asp  
384

Nostoc.ST25.txt

115	120	125	
ttt cat gat ggt aag aga aca aac gct att ttc tgg tat ctc cat ttc Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe 130 135 140			432
atg ata gaa tac tcc agt tgg caa cag tta ata gta cta act atc cta Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu 145 150 155 160			480
ttt aat tta gct aaa tac gtt ttg cac atc cat caa ata aat ctc atc Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile 165 170 175			528
tta ttt tgg agt att cct cca att tta agt tcc att caa ctg ttt tat Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr 180 185 190			576
ttc gga aca ttt ttg cct cat cga gaa ccc aag aaa gga tat gtt tat Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr 195 200 205			624
cat tgc agc caa aca ata aaa ttg cca act ttt ttg tca ttt atc His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile 210 215 220			672
gct tgc tac cac ttt ggt tat cat gaa gaa cat cat gag tat ccc cat Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His 225 230 235 240			720
gta cct tgg tgg caa ctt cca tct gta tat aag cag aga gta ttc aac Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn 245 250 255			768
aat tca gta acc aat tcg taa Asn Ser Val Thr Asn Ser 260			789

<210> 10

<211> 262

<212> PRT

3> Künstliche Sequenz

<400> 10

Met Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln  
1 5 10 15

Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val  
20 25 30

Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn  
35 40 45

Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln  
50 55 60

Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His  
Seite 12

Nostoc.ST25.txt  
75

65

70

80

Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser  
85 90 95

Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys  
100 105 110

Asn His Cys Leu His His Arg His Pro Ala Ser Asp Leu Asp Pro Asp  
115 120 125

Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe  
130 135 140

Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu  
145 150 155 160

Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile  
165 170 175

Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr  
180 185 190

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr  
195 200 205

Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile  
210 215 220

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
225 230 235 240

Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn  
245 250 255

Asn Ser Val Thr Asn Ser  
260

<210> 11

<211> 762

<212> DNA

<213> Künstliche Sequenz

<220>

<221> CDS

<222> (1)..(762)

<223>

Nostoc.ST25.txt

<400> 11	48
atg atc cag tta gaa caa cca ctc agt cat caa gca aaa ctg act cca Met Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro 1 5 10 15	
gta ctg aga agt aaa tct cag ttt aag ggg ctt ttc att gct att gtc Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val 20 25 30	96
att gtt agc gca tgg gtc att agc ctg agt tta tta ctt tcc ctt gac Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp 35 40 45	144
atc tca aag att cat aag tgg atg tta ttg cct gtt ata cta tgg caa Ile Ser Lys Ile His Lys Trp Met Leu Leu Pro Val Ile Leu Trp Gln 50 55 60	192
aca ttt tta tat acg gga tta ttt att aca tct cat gat gcc atg cat Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His 70 75 80	240
gta gta ttt ccc caa aac acc aag att aat cat ttg att gga aca Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr 85 90 95	288
ttg acc cta tcc ctt tat ggt ctt tta cca tat caa aaa cta ttg aaa Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys 100 105 110	336
aaa cat tgg tta cac cac cac aat cca gca agc tca ata gac ccg gat Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp 115 120 125	384
ttt cac aat ggt aaa cac caa agt ttc ttt gct tgg tat ttt cat ttt Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe 130 135 140	432
atg aaa ggt tac tgg agt tgg ggg caa ata att gcg ttg act att att Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile 145 150 155 160	480
t aac ttt gct aaa tac ata ctc cat atc cca agt gat aat cta act Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr 165 170 175	528
tac ttt tgg gtg cta ccc tcg ctt tta agt tca tta caa tta ttc tat Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr 180 185 190	576
ttt ggt act ttt tta ccc cat agt gaa cca ata ggg ggt tat gtt cag Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln 195 200 205	624
cct cat tgt gcc caa aca att agc cgt cct att tgg tgg tca ttt atc Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile 210 215 220	672
acg tgc tat cat ttt ggc tac cac gag gaa cat cac gaa tat cct cat Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His 225 230 235 240	720
att tct tgg tgg cag tta cca gaa att tac aaa gca aaa tag Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys 245 250	762

Nostoc.ST25.txt

<210> 12

<211> 253

<212> PRT

<213> Künstliche Sequenz

<400> 12

Met Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro  
1 5 10 15

Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val  
20 25 30

Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp  
35 40 45

Ile Ser Lys Ile His Lys Trp Met Leu Leu Pro Val Ile Leu Trp Gln  
50 55 60

Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His  
65 70 75 80

Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr  
85 90 95

Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys  
100 105 110

Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp  
115 120 125

His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe  
130 135 140

Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile  
145 150 155 160

Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr  
165 170 175

Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr  
180 185 190

Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln  
195 200 205

Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile  
210 215 220

Nostoc.ST25.txt

Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
225 230 235 240

Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys  
245 250

<210> 13

<211> 762

<212> DNA

<213> künstliche Sequenz

<220>

<221> CDS

<222> (1)..(762)

<223>

<400> 13	atg atc cag tta gaa caa cca ctc agt cat caa gca aaa ctg act cca	48
Met Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro	1 5 10 15	
gta ctg aga agt aaa tct cag ttt aag ggg ctt ttc att gct att gtc	96	
Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val	20 25 30	
att gtt agc gca tgg gtc att agc ctg agt tta tta ctt tcc ctt gac	144	
Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Leu Ser Leu Asp	35 40 45	
atc tca aag cta aaa ttt tgg atg tta ttg cct gtt ata cta tgg caa	192	
Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln	50 55 60	
aaa ttt tta tat acg gga tta ttt att aca tct cat gat gcc atg cat	240	
Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His	65 70 75 80	
ggc gta gta ttt ccc caa aac acc aag att aat cat ttg att gga aca	288	
Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr	85 90 95	
ttg acc cta tcc ctt tat ggt ctt tta cca tat caa aaa cta ttg aaa	336	
Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys	100 105 110	
aaa cat tgg tta cac cac cac aat cca gca agc gat tta gac ccg gat	384	
Lys His Trp Leu His His His Asn Pro Ala Ser Asp Leu Asp Pro Asp	115 120 125	
ttt cac aat ggt aaa cac caa agt ttc ttt gct tgg tat ttt cat ttt	432	
Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe	130 135 140	
atg aaa ggt tac tgg agt tgg ggg caa ata att gcg ttg act att att	480	
Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile		



Nostoc.ST25.txt  
155

145	150	160	
tat aac ttt gct aaa tac ata ctc cat atc cca agt gat aat cta act			528
Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr			
	165	170	
tac ttt tgg gtg cta ccc tcg ctt tta agt tca tta caa tta ttc tat			576
Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr			
	180	185	
ttt ggt act ttt tta ccc cat agt gaa cca ata ggg ggt tat gtt cag			624
Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln			
	195	200	
cct cat tgt gcc caa aca att agc cgt cct att tgg tgg tca ttt atc			672
Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile			
	210	215	
acg tgc tat cat ttt ggc tac cac gag gaa cat cac gaa tat cct cat			720
Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His			
	225	230	
tct tgg tgg cag tta cca gaa att tac aaa gca aaa tag			762
Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys			
	245	250	

<210> 14

<211> 253

<212> PRT

<213> Künstliche Sequenz

<400> 14

Met Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro  
1 5 10 15

Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val  
20 25 30

Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp  
35 40 45

Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln  
50 55 60

Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His  
65 70 75 80

Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr  
85 90 95

Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys  
100 105 110

Lys His Trp Leu His His His Asn Pro Ala Ser Asp Leu Asp Pro Asp  
seite 17

Nostoc.ST25.txt

115

120

125

Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe  
130 135 140

Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile  
145 150 155 160

Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr  
165 170 175

Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr  
180 185 190

Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln  
195 200 205

His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile  
210 215 220

Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
225 230 235 240

Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys  
245 250

<210> 15

<211> 1608

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (3)...(971)

<223>

<400> 15

ct aca ttt cac aag ccc gtg agc ggt gca agc gct ctg ccc cac atc 47  
Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile  
1 5 10 15

ggc cca cct cct cat ctc cat cgg tca ttt gct gct acc acg atg ctg 95  
Gly Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu  
20 25 30

tcg aag ctg cag tca atc agc gtc aag gcc cgc cgc gtt gaa cta gcc 143  
Ser Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala  
35 40 45

Nostoc.ST25.txt

cgc gac atc acg cgg ccc aaa gtc tgc ctg cat gct cag cgg tgc tcg Arg Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser 50 55 60	191
tta gtt cgg ctg cga gtg gca gca cca cag aca gag gag gcg ctg gga Leu Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly 65 70 75	239
acc gtg cag gct gcc ggc gcg ggc gat gag cac agc gcc gat gta gca Thr Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala 80 85 90 95	287
ctc cag cag ctt gac cgg gct atc gca gag cgt cgt gcc cgg cgc aaa Leu Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys 100 105 110	335
cgg gag cag ctg tca tac cag gct gcc gcc att gca gca tca att ggc Arg Glu Gln Leu Ser Tyr Gln Ala Ala Ile Ala Ala Ser Ile Gly 115 120 125	383
gtg tca ggc att gcc atc ttc gcc acc tac ctg aga ttt gcc atg cac Val Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His 130 135 140	431
gag acc gtg ggc ggc gca gtg cca tgg ggt gaa gtg gct ggc act ctc Met Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu 145 150 155	479
ctc ttg gtg gtt ggt ggc gcg ctc ggc atg gag atg tat gcc cgc tat Leu Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr 160 165 170 175	527
gca cac aaa gcc atc tgg cat gag tgc cct ctg ggc tgg ctg ctg cac Ala His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His 180 185 190	575
aag agc cac cac aca cct cgc act gga ccc ttt gaa gcc aac gac ttg Lys Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu 195 200 205	623
ttt gca atc atc aat gga ctg ccc gcc atg ctc ctg tgt acc ttt ggc Phe Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly 210 215 220	671
ctc tgg ctg ccc aac gtc ctg ggg gcg gcc tgc ttt gga gcg ggg ctg Val Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu 225 230 235	719
ggc atc acg cta tac ggc atg gca tat atg ttt gta cac gat ggc ctg Gly Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu 240 245 250 255	767
gtg cac agg cgc ttt ccc acc ggg ccc atc gct ggc ctg ccc tac atg Val His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met 260 265 270	815
aag cgc ctg aca gtg gcc cac cag cta cac cac agc ggc aag tac ggt Lys Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly 275 280 285	863
ggc gcg ccc tgg ggt atg ttc ttg ggt cca cag gag ctg cag cac att Gly Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile 290 295 300	911
cca ggt gcg gcg gag gag gtg gag cga ctg gtc ctg gaa ctg gac tgg Pro Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp 305 310 315	959

Nostoc.ST25.txt

tcc aag cgg tag ggtgCGgaac caggcagct ggtttcacac ctcatgcctg 1011  
Ser Lys Arg  
320

tgataaggtg tggctagagc gatgcgtgtg agacgggtat gtcacggtcg actggtctga 1071  
tggccaatgg catcggccat gtctgggtcat cacgggctgg ttgcctgggt gaaggtgatg 1131  
cacatcatca tgtgcggttg gaggggctgg cacagtgtgg gctgaactgg agcagttgtc 1191  
caggctggcg ttgaatcagt gagggtttgt gattggcgggt tgtgaagcaa tgactccgcc 1251  
catattctat ttgtgggagc tgagatgatg gcatgcttgg gatgtgcatg gatcatggta 1311  
gtgcagcaaa ctatattcac ctagggctgt tggtaggatc aggtgaggcc ttgcacattg 1371  
catgatgtac tcgtcatggt gtgttgggtga gaggatggat gtggatggat gtgtattctc 1431  
agacgtagac cttgactgga ggcttgatcg agagagtggg ccgtattctt tgagagggga 1491  
ggctcgtgcc agaaatgggtg agtggatgac tgtgacgctg tacattgcag gcaggtgaga 1551  
ctgtct cgattgtaaa atacattcag atgcaaaaaa aaaaaaaaaa aaaaaa 1608

<210> 16

<211> 322

<212> PRT

<213> Haematococcus pluvialis

<400> 16

Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile Gly  
1 5 10 15

Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu Ser  
20 25 30

Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala Arg  
35 40 45

Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser Leu  
50 55 60

Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly Thr  
65 70 75 80

Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala Leu  
85 90 95

Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys Arg  
100 105 110

Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly Val  
115 120 125

Nostoc.ST25.txt

Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His Met  
130 135 140

Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu Leu  
145 150 155 160

Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr Ala  
165 170 175

His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His Lys  
180 185 190

Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu Phe  
195 200 205

Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly Phe  
210 215 220

Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu Gly  
225 230 235 240

Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu Val  
245 250 255

His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met Lys  
260 265 270

Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly Gly  
275 280 285

Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile Pro  
290 295 300

Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp Ser  
305 310 315 320

Lys Arg

<210> 17

<211> 1650

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> CDS

<222> (112)..(1614)

Nostoc.ST25.txt

<223>

<400> 17  
ggcaccgagga aactttttctc tcttcactag ctgtttacat gcttgaaatt tcaagatttt 60  
aggaccccat ttgaagtttt cttgaaacaa atattaccct gttggaaaaa g atg gat 117  
Met Asp  
1  
act ttg ttg aaa acc cca aat aac ctt gaa ttt ctg aac cca cat cat 165  
Thr Leu Leu Lys Thr Pro Asn Asn Leu Glu Phe Leu Asn Pro His His  
5 10 15  
ggt ttt gct gtt aaa gct agt acc ttt aga tct gag aag cat cat aat 213  
Gly Phe Ala Val Lys Ala Ser Thr Phe Arg Ser Glu Lys His His Asn  
20 25 30  
ttt ggt tct agg aag ttt tgt gaa act ttg ggt aga agt gtt tgt gtt 261  
Gly Ser Arg Lys Phe Cys Glu Thr Leu Gly Arg Ser Val Cys Val  
40 45 50  
aag ggt agt agt agt gct ctt tta gag ctt gta cct gag acc aaa aag 309  
Lys Gly Ser Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr Lys Lys  
55 60 65  
gag aat ctt gat ttt gag ctt cct atg tat gac cct tca aaa ggg gtt 357  
Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys Gly Val  
70 75 80  
gtt gtg gat ctt gct gtg gtt ggt ggt ggc cct gca gga ctt gct gtt 405  
Val Val Asp Leu Ala Val Val Gly Gly Gly Pro Ala Gly Leu Ala Val  
85 90 95  
gca cag caa gtt tct gaa gca gga ctc tct gtt tgt tca att gat ccg 453  
Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp Pro  
100 105 110  
aat cct aaa ttg ata tgg cct aat aac tat ggt gtt tgg gtg gat gaa 501  
Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu  
115 120 125 130  
gag gct atg gac ttg tta gat tgt cta gat gct acc tgg tct ggt 549  
Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp Ser Gly  
135 140 145  
gca gca gtg tac att gat gat aat acg gct aaa gat ctt cat aga cct 597  
Ala Ala Val Tyr Ile Asp Asp Asn Thr Ala Lys Asp Leu His Arg Pro  
150 155 160  
tat gga agg gtt aac cgg aaa cag ctg aaa tcg aaa atg atg cag aaa 645  
Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met Gln Lys  
165 170 175  
tgt ata atg aat ggt gtt aaa ttc cac caa gcc aaa gtt ata aag gtg 693  
Cys Ile Met Asn Gly Val Lys Phe His Gln Ala Lys Val Ile Lys Val  
180 185 190  
att cat gag gaa tcg aaa tcc atg ttg ata tgc aat gat ggt att act 741  
Ile His Glu Glu Ser Lys Ser Met Leu Ile Cys Asn Asp Gly Ile Thr  
195 200 205 210  
att cag gca acg gtg gtg ctc gat gca act ggc ttc tct aga tct ctt 789  
Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg Ser Leu  
215 220 225

Nostoc.ST25.txt

ggt cag tat gat aag cct tat aac ccc ggg tat caa gtt gct tat ggc Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala Tyr Gly 230 235 240	837
att ttg gct gaa gtg gaa gag cac ccc ttt gat gta aac aag atg gtt Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Val Asn Lys Met Val 245 250 255	885
ttc atg gat tgg cga gat tct cat ttg aag aac aat act gat ctc aag Phe Met Asp Trp Arg Asp Ser His Leu Lys Asn Asn Thr Asp Leu Lys 260 265 270	933
gag aga aat agt aga ata cca act ttt ctt tat gca atg cca ttt tca Glu Arg Asn Ser Arg Ile Pro Thr Phe Leu Tyr Ala Met Pro Phe Ser 275 280 285 290	981
tcc aac agg ata ttt ctt gaa gaa aca tca ctc gta gct cgt cct ggc Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg Pro Gly 295 300 305	1029
ttg cgt ata gat gat att caa gaa cga atg gtg gct cgt tta aac cat Arg Ile Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu Asn His 310 315 320	1077
ttg ggg ata aaa gtg aag agc att gaa gaa gat gaa cat tgt cta ata Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys Leu Ile 325 330 335	1125
cca atg ggt ggt cca ctt cca gta tta cct cag aga gtc gtt gga atc Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val Gly Ile 340 345 350	1173
ggt ggt aca gct ggc atg gtt cat cca tcc acc ggt tat atg gtg gca Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val Ala 355 360 365 370	1221
agg aca cta gct gcg gct cct gtt gtt gcc aat gcc ata att caa tac Arg Thr Leu Ala Ala Pro Val Val Ala Asn Ala Ile Ile Gln Tyr 375 380 385	1269
ctc ggt tct gaa aga agt cat tcg ggt aat gaa tta tcc aca gct gtt Leu Gly Ser Glu Arg Ser His Ser Gly Asn Glu Leu Ser Thr Ala Val 390 395 400	1317
agg aaa gat ttg tgg cct ata gag agg aga cgt caa aga gag ttc ttc Lys Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln Arg Glu Phe Phe 405 410 415	1365
tgc ttc ggt atg gat att ctt ctg aag ctt gat tta cct gct aca aga Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala Thr Arg 420 425 430	1413
agg ttc ttt gat gca ttc ttt gac tta gaa cct cgt tat tgg cat ggc Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp His Gly 435 440 445 450	1461
ttc tta tcg tct cga ttg ttt cta cct gaa ctc ata gtt ttt ggg ctg Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Ile Val Phe Gly Leu 455 460 465	1509
tct cta ttc tct cat gct tca aat act tct aga ttt gag ata atg aca Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Phe Glu Ile Met Thr 470 475 480	1557
aag gga act gtt cca tta gta aat atg atc aac aat ttg tta cag gat Lys Gly Thr Val Pro Leu Val Asn Met Ile Asn Asn Leu Leu Gln Asp 485 490 495	1605

aaa gaa tga atccgagtaa ttcggaatct tgtccaatct cgtgcc  
 Lys Glu  
 500

<210> 18

<211> 500

<212> PRT

<213> Lycopersicon esculentum

<400> 18

Met Asp Thr Leu Leu Lys Thr Pro Asn Asn Leu Glu Phe Leu Asn Pro  
 1 5 10 15

His Gly Phe Ala Val Lys Ala Ser Thr Phe Arg Ser Glu Lys His  
 20 25 30

His Asn Phe Gly Ser Arg Lys Phe Cys Glu Thr Leu Gly Arg Ser Val  
 35 40 45

Cys Val Lys Gly Ser Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr  
 50 55 60

Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys  
 65 70 75 80

Gly Val Val Val Asp Leu Ala Val Val Gly Gly Gly Pro Ala Gly Leu  
 85 90 95

Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile  
 100 105 110

Pro Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val  
 115 120 125

Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp  
 130 135 140

Ser Gly Ala Ala Val Tyr Ile Asp Asp Asn Thr Ala Lys Asp Leu His  
 145 150 155 160

Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met  
 165 170 175

Gln Lys Cys Ile Met Asn Gly Val Lys Phe His Gln Ala Lys Val Ile  
 180 185 190

Lys Val Ile His Glu Glu Ser Lys Ser Met Leu Ile Cys Asn Asp Gly  
 195 200 205



Nostoc.ST25.txt

Ile Thr Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg  
210 215 220

Ser Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala  
225 230 235 240

Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Val Asn Lys  
245 250 255

Met Val Phe Met Asp Trp Arg Asp Ser His Leu Lys Asn Asn Thr Asp  
260 265 270

Leu Lys Glu Arg Asn Ser Arg Ile Pro Thr Phe Leu Tyr Ala Met Pro  
275 280 285

Ser Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg  
290 295 300

Pro Gly Leu Arg Ile Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu  
305 310 315 320

Asn His Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys  
325 330 335

Leu Ile Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val  
340 345 350

Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met  
355 360 365

Val Ala Arg Thr Leu Ala Ala Ala Pro Val Val Ala Asn Ala Ile Ile  
370 375 380

Tyr Leu Gly Ser Glu Arg Ser His Ser Gly Asn Glu Leu Ser Thr  
390 395 400

Ala Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln Arg Glu  
405 410 415

Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala  
420 425 430

Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp  
435 440 445

His Gly Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Ile Val Phe  
450 455 460

Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Phe Glu Ile  
465 470 475 480

Nostoc.ST25.txt

Met Thr Lys Gly Thr Val Pro Leu Val Asn Met Ile Asn Asn Leu Leu  
 485 490 495

Gln Asp Lys Glu  
 500

<210> 19

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<221> primer\_bind

<222> (1)..(33)

<223>

<400> 19  
 gcatgctcta gaccttataa agatattttg tga

33

<210> 20

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<221> primer\_bind

<222> (1)..(33)

<223>

<400> 20  
 gcatgcatct agaaatgggt cagtgtcaac cat

33

<210> 21

<211> 805

<212> DNA

<213> Nostoc sp. Strain PCC7120

<220>

Nostoc.ST25.txt

<221> variation

<222> (1)..(805)

<223>

<400> 21

gcatgcatct agaaatgggt cagtgtcaac catcatctct gcattcagaa aaactgggtgt	60
tattgtcatc gacaatcaga gatgataaaa atattaataa gggatatatt attgcctgct	120
ttatcttatt tttatgggca attagtttaa tcttattact ctcaatagat acatccataa	180
ttcataagag cttattaggt atagccatgc tttggcagac cttcttatat acaggtttat	240
ttattactgc tcatgatgcc atgcacggcg tagtttatcc caaaaatccc agaataaata	300
atctttatagg taagctcact ctaatcttgt atggactact cccttataaa gattttattga	360
gtacattg gttacaccac ggacatcctg gtactgattt agaccctgat tattacaatg	420
gtcatcccca aaacttcttt ctttgggtatc tacattttat gaagtcttat tggcgatgga	480
cgcaaatttt cggattagtg atgatttttc atggacttaa aaatctggtg catataccag	540
aaaataattt aattatattt tggatgatac cttctatttt aagttcagta caactatttt	600
atcttggtac atctttgcct cataaaaagc tagaagggtg ttatactaac cccattgtg	660
cgcgcagtat cccattacct cttttttggt cttttgttac ttgttatcac ttcggctacc	720
acaaggaaca tcacgaatac cctcaacttc cttggtggaa attacctgaa gctcacaaaa	780
tatctttata aggtctagag catgc	805

<210> 22

<211> 24

<212> DNA

<213> Künstliche Sequenz

<220>

<221> primer\_bind

<222> (1)..(24)

<223>

<400> 22

aggtaccgca cggctctgcca atcc

24

<210> 23

<211> 26

<212> DNA

<213> Künstliche Sequenz

<220>

<221> primer\_bind

<222> (1)..(26)

<223>

<400> 23  
aagcttgacc tgattatcag cacggt

26

<210> 24

<211> 4624

<212> DNA

<213> Erwinia uredovora

<220>

<221> CDS

<222> (128)..(1267)

<223>

<220>

<221> CDS

<222> (1288)..(2766)

<223>

<220>

<221> CDS

<222> (2802)..(3689)

<223>

<220>

<221> iDNA

<222> (3631)..(4158)

<223>

Nostoc.ST25.txt

<400> 24  
gtcgactttc agcagcgcat ggcgaaaatc cagacagccc ttcgtttggc agggggcacc 60  
atggccgctg ccgatatcat tgagcagggtt atgtgcaccg gtcagcctgt cttaagtggg 120  
agcggct atg caa ccg cat tat gat ctg att ctc gtg ggg gct gga ctc 169  
Met Gln Pro His Tyr Asp Leu Ile Leu Val Gly Ala Gly Leu 10  
gcg aat ggc ctt atc gcc ctg cgt ctt cag cag cag caa cct gat atg 217  
Ala Asn Gly Leu Ile Ala Leu Arg Leu Gln Gln Gln Gln Pro Asp Met 30  
15 20 25 30  
cgt att ttg ctt atc gac gcc gca ccc cag gcg ggc ggg aat cat acg 265  
Arg Ile Leu Leu Ile Asp Ala Ala Pro Gln Ala Gly Gly Asn His Thr 45  
35 40 45  
tgg tca ttt cac cac gat gat ttg act gag agc caa cat cgt tgg ata 313  
Trp Ser Phe His His Asp Asp Leu Thr Glu Ser Gln His Arg Trp Ile 60  
50 55 60  
ccg ctg gtg gtt cat cac tgg ccc gac tat cag gta cgc ttt ccc 361  
Pro Leu Val Val His His Trp Pro Asp Tyr Gln Val Arg Phe Pro 75  
65 70 75  
aca cgc cgt cgt aag ctg aac agc ggc tac ttt tgt att act tct cag 409  
Thr Arg Arg Arg Lys Leu Asn Ser Gly Tyr Phe Cys Ile Thr Ser Gln 90  
80 85 90  
cgt ttc gct gag gtt tta cag cga cag ttt ggc ccg cac ttg tgg atg 457  
Arg Phe Ala Glu Val Leu Gln Arg Gln Phe Gly Pro His Leu Trp Met 110  
95 100 105 110  
gat acc gcg gtc gca gag gtt aat gcg gaa tct gtt cgg ttg aaa aag 505  
Asp Thr Ala Val Ala Glu Val Asn Ala Glu Ser Val Arg Leu Lys Lys 125  
115 120 125  
ggg cag gtt atc ggt gcc cgc gcg gtg att gac ggg cgg ggt tat gcg 553  
Gly Gln Val Ile Gly Ala Arg Ala Val Ile Asp Gly Arg Gly Tyr Ala 140  
130 135 140  
gca aat tca gca ctg agc gtg ggc ttc cag gcg ttt att ggc cag gaa 601  
Ala Asn Ser Ala Leu Ser Val Gly Phe Gln Ala Phe Ile Gly Gln Glu 155  
145 150 155  
cga ttg agc cac ccg cat ggt tta tcg tct ccc att atc atg gat 649  
Trp Arg Leu Ser His Pro His Gly Leu Ser Ser Pro Ile Ile Met Asp 170  
160 165 170  
gcc acg gtc gat cag caa aat ggt tat cgc ttc gtg tac agc ctg ccg 697  
Ala Thr Val Asp Gln Gln Asn Gly Tyr Arg Phe Val Tyr Ser Leu Pro 190  
175 180 185 190  
ctc tcg ccg acc aga ttg tta att gaa gac acg cac tat att gat aat 745  
Leu Ser Pro Thr Arg Leu Leu Ile Glu Asp Thr His Tyr Ile Asp Asn 205  
195 200 205  
gcg aca tta gat cct gaa tgc gcg cgg caa aat att tgc gac tat gcc 793  
Ala Thr Leu Asp Pro Glu Cys Ala Arg Gln Asn Ile Cys Asp Tyr Ala 220  
210 215 220  
gcg caa cag ggt tgg cag ctt cag aca ctg ctg cga gaa gaa cag ggc 841  
Ala Gln Gln Gly Trp Gln Leu Gln Thr Leu Leu Arg Glu Glu Gln Gly 235  
225 230 235  
gcc tta ccc att act ctg tcg ggc aat gcc gac gca ttc tgg cag cag 889  
Ala Leu Pro Ile Thr Leu Ser Gly Asn Ala Asp Ala Phe Trp Gln Gln  
240 245 250

Nostoc.ST25.txt  
250

240	245		937
cgc ccc ctg gcc tgt agt gga tta cgt gcc ggt ctg ttc cat cct acc	Arg Pro Leu Ala Cys Ser Gly Leu Arg Ala Gly Leu Phe His Pro Thr	255 260 265 270	
acc ggc tat tca ctg ccg ctg gcg gtt gcc gtg gcc gac cgc ctg agt	Thr Gly Tyr Ser Leu Pro Leu Ala Val Ala Val Ala Asp Arg Leu Ser	275 280 285	985
gca ctt gat gtc ttt acg tcg gcc tca att cac cat gcc att acg cat	Ala Leu Asp Val Phe Thr Ser Ala Ser Ile His His Ala Ile Thr His	290 295 300	1033
ttt gcc cgc gag cgc tgg cag cag cag ggc ttt ttc cgc atg ctg aat	Phe Ala Arg Glu Arg Trp Gln Gln Gln Gly Phe Phe Arg Met Leu Asn	305 310 315	1081
cgc atg ctg ttt tta gcc gga ccc gcc gat tca cgc tgg cgg gtt atg	Arg Met Leu Phe Leu Ala Gly Pro Ala Asp Ser Arg Trp Arg Val Met	320 325 330	1129
cgt ttt tat ggt tta cct gaa gat tta att gcc cgt ttt tat gcg	Arg Phe Tyr Gly Leu Pro Glu Asp Leu Ile Ala Arg Phe Tyr Ala	340 345 350	1177
gga aaa ctc acg ctg acc gat cgg cta cgt att ctg agc ggc aag ccg	Gly Lys Leu Thr Leu Thr Asp Arg Leu Arg Ile Leu Ser Gly Lys Pro	355 360 365	1225
cct gtt ccg gta tta gca gca ttg caa gcc att atg acg act	Pro Val Pro Val Leu Ala Ala Leu Gln Ala Ile Met Thr Thr	370 375 380	1267
catcgttaaa gagcgactac atg aaa cca act acg gta att ggt gca ggc ttc	Met Lys Pro Thr Thr Val Ile Gly Ala Gly Phe	385 390	1320
ggt ggc ctg gca ctg gca att cgt cta caa gct gcg ggg atc ccc gtc	Gly Gly Leu Ala Leu Ala Ile Arg Leu Gln Ala Ala Gly Ile Pro Val	395 400 405	1368
tta ctg ctt gaa caa cgt gat aaa ccc ggc ggt cgg gct tat gtc tac	Leu Leu Leu Glu Gln Arg Asp Lys Pro Gly Gly Arg Ala Tyr Val Tyr	410 415 420	1416
gat cag ggg ttt acc ttt gat gca ggc ccg acg gtt atc acc gat	Glu Asp Gln Gly Phe Thr Phe Asp Ala Gly Pro Thr Val Ile Thr Asp	425 430 435	1464
ccc agt gcc att gaa gaa ctg ttt gca ctg gca gga aaa cag tta aaa	Pro Ser Ala Ile Glu Glu Leu Phe Ala Leu Ala Gly Lys Gln Leu Lys	440 445 450 455	1512
gag tat gtc gaa ctg ctg ccg gtt acg ccg ttt tac cgc ctg tgt tgg	Glu Tyr Val Glu Leu Leu Pro Val Thr Pro Phe Tyr Arg Leu Cys Trp	460 465 470	1560
gag tca ggg aag gtc ttt aat tac gat aac gat caa acc cgg ctc gaa	Glu Ser Gly Lys Val Phe Asn Tyr Asp Asn Asp Gln Thr Arg Leu Glu	475 480 485	1608
gcg cag att cag cag ttt aat ccc cgc gat gtc gaa ggt tat cgt cag	Ala Gln Ile Gln Gln Phe Asn Pro Arg Asp Val Glu Gly Tyr Arg Gln	490 495 500	1656
ttt ctg gac tat tca cgc gcg gtg ttt aaa gaa ggc tat cta aag ctc	Phe Leu Asp Tyr Ser Arg Ala Val Phe Lys Glu Gly Tyr Leu Lys Leu		1704

Nostoc.ST25.txt  
515

505	510		1752
ggt act gtc cct ttt tta tcg ttc aga gac atg ctt cgc gcc gca cct Gly Thr Val Pro Phe Leu Ser Phe Arg Asp Met Leu Arg Ala Ala Pro 520 525 530 535			
caa ctg gcg aaa ctg cag gca tgg aga agc gtt tac agt aag gtt gcc Gln Leu Ala Lys Leu Gln Ala Trp Arg Ser Val Tyr Ser Lys Val Ala 540 545 550			1800
agt tac atc gaa gat gaa cat ctg cgc cag gcg ttt tct ttc cac tcg Ser Tyr Ile Glu Asp Glu His Leu Arg Gln Ala Phe Ser Phe His Ser 555 560 565			1848
ctg ttg gtg ggc ggc aat ccc ttc gcc acc tca tcc att tat acg ttg Leu Leu Val Gly Gly Asn Pro Phe Ala Thr Ser Ser Ile Tyr Thr Leu 570 575 580			1896
ata cac gcg ctg gag cgt gag tgg ggc gtc tgg ttt ccg cgt ggc ggc Ile His Ala Leu Glu Arg Glu Trp Gly Val Trp Phe Pro Arg Gly Gly 585 590 595			1944
ggc gca tta gtt cag ggg atg ata aag ctg ttt cag gat ctg ggt Gly Ala Leu Val Gln Gly Met Ile Lys Leu Phe Gln Asp Leu Gly 600 605 610 615			1992
ggc gaa gtc gtg tta aac gcc aga gtc agc cat atg gaa acg aca gga Gly Glu Val Val Leu Asn Ala Arg Val Ser His Met Glu Thr Thr Gly 620 625 630			2040
aac aag att gaa gcc gtg cat tta gag gac ggt cgc agg ttc ctg acg Asn Lys Ile Glu Ala Val His Leu Glu Asp Gly Arg Arg Phe Leu Thr 635 640 645			2088
caa gcc gtc gcg tca aat gca gat gtg gtt cat acc tat cgc gac ctg Gln Ala Val Ala Ser Asn Ala Asp Val Val His Thr Tyr Arg Asp Leu 650 655 660			2136
tta agc cag cac cct gcc gcg gtt aag cag tcc aac aaa ctg cag act Leu Ser Gln His Pro Ala Ala Val Lys Gln Ser Asn Lys Leu Gln Thr 665 670 675			2184
aag cgc atg agt aac tct ctg ttt gtg ctc tat ttt ggt ttg aat cac Lys Arg Met Ser Asn Ser Leu Phe Val Leu Tyr Phe Gly Leu Asn His 680 685 690 695			2232
ct cat gat cag ctc gcg cat cac acg gtt tgt ttc ggc ccg cgt tac His His Asp Gln Leu Ala His His Thr Val Cys Phe Gly Pro Arg Tyr 700 705 710			2280
cgc gag ctg att gac gaa att ttt aat cat gat ggc ctc gca gag gac Arg Glu Leu Ile Asp Glu Ile Phe Asn His Asp Gly Leu Ala Glu Asp 715 720 725			2328
ttc tca ctt tat ctg cac gcg ccc tgt gtc acg gat tcg tca ctg gcg Phe Ser Leu Tyr Leu His Ala Pro Cys Val Thr Asp Ser Ser Leu Ala 730 735 740			2376
cct gaa ggt tgc ggc agt tac tat gtg ttg gcg ccg gtg ccg cat tta Pro Glu Gly Cys Gly Ser Tyr Tyr Val Leu Ala Pro Val Pro His Leu 745 750 755			2424
ggc acc gcg aac ctc gac tgg acg gtt gag ggg cca aaa cta cgc gac Gly Thr Ala Asn Leu Asp Trp Thr Val Glu Gly Pro Lys Leu Arg Asp 760 765 770 775			2472
cgt att ttt gcg tac ctt gag cag cat tac atg cct ggc tta cgg agt Arg Ile Phe Ala Tyr Leu Glu Gln His Tyr Met Pro Gly Leu Arg Ser 780 785 790 795			2520

Nostoc.ST25.txt

780										785					790					
cag	ctg	gtc	acg	cac	cgg	atg	ttt	acg	ccg	ttt	gat	ttt	cgc	gac	cag	2568				
Gln	Leu	Val	Thr	His	Arg	Met	Phe	Thr	Pro	Phe	Asp	Phe	Arg	Asp	Gln					
			795					800					805							
ctt	aat	gcc	tat	cat	ggc	tca	gcc	ttt	tct	gtg	gag	ccc	ggt	ctt	acc	2616				
Leu	Asn	Ala	Tyr	His	Gly	Ser	Ala	Phe	Ser	Val	Glu	Pro	Val	Leu	Thr					
		810					815					820								
cag	agc	gcc	tgg	ttt	cgg	ccg	cat	aac	cgc	gat	aaa	acc	att	act	aat	2664				
Gln	Ser	Ala	Trp	Phe	Arg	Pro	His	Asn	Arg	Asp	Lys	Thr	Ile	Thr	Asn					
	825					830					835									
ctc	tac	ctg	gtc	ggc	gca	ggc	acg	cat	ccc	ggc	gca	ggc	att	cct	ggc	2712				
Leu	Tyr	Leu	Val	Gly	Ala	Gly	Thr	His	Pro	Gly	Ala	Gly	Ile	Pro	Gly					
840					845					850					855					
gtc	atc	ggc	tcg	gca	aaa	gcg	aca	gca	ggt	ttg	atg	ctg	gag	gat	ctg	2760				
Val	Ile	Gly	Ser	Ala	Lys	Ala	Thr	Ala	Gly	Leu	Met	Leu	Glu	Asp	Leu					
				860					865					870						
tga	ataatccg	tc	gttactcaat	catgcg	gtcg	aaacg	atg	gca	ggt	ggc						2813				
							Met	Ala	Val	Gly					875					
tcg	aaa	agt	ttt	gcg	aca	gcc	tca	aag	tta	ttt	gat	gca	aaa	acc	cgg	2861				
Ser	Lys	Ser	Phe	Ala	Thr	Ala	Ser	Lys	Leu	Phe	Asp	Ala	Lys	Thr	Arg					
			880					885					890							
cgc	agc	gta	ctg	atg	ctc	tac	gcc	tgg	tgc	cgc	cat	tgt	gac	gat	gtt	2909				
Arg	Ser	Val	Leu	Met	Leu	Tyr	Ala	Trp	Cys	Arg	His	Cys	Asp	Asp	Val					
		895					900					905								
att	gac	gat	cag	acg	ctg	ggc	ttt	cag	gcc	cgg	cag	cct	gcc	tta	caa	2957				
Ile	Asp	Asp	Gln	Thr	Leu	Gly	Phe	Gln	Ala	Arg	Gln	Pro	Ala	Leu	Gln					
	910					915					920									
acg	ccc	gaa	caa	cgt	ctg	atg	caa	ctt	gag	atg	aaa	acg	cgc	cag	gcc	3005				
Thr	Pro	Glu	Gln	Arg	Leu	Met	Gln	Leu	Glu	Met	Lys	Thr	Arg	Gln	Ala					
	925				930					935					940					
tat	gca	gga	tcg	cag	atg	cac	gaa	ccg	gcg	ttt	gcg	gct	ttt	cag	gaa	3053				
Thr	Ala	Gly	Ser	Gln	Met	His	Glu	Pro	Ala	Phe	Ala	Ala	Phe	Gln	Glu					
				945					950					955						
gtg	gct	atg	gct	cat	gat	atc	gcc	ccg	gct	tac	gcg	ttt	gat	cat	ctg	3101				
Val	Ala	Met	Ala	His	Asp	Ile	Ala	Pro	Ala	Tyr	Ala	Phe	Asp	His	Leu					
			960					965					970							
gaa	ggc	ttc	gcc	atg	gat	gta	cgc	gaa	gcg	caa	tac	agc	caa	ctg	gat	3149				
Glu	Gly	Phe	Ala	Met	Asp	Val	Arg	Glu	Ala	Gln	Tyr	Ser	Gln	Leu	Asp					
		975					980					985								
gat	acg	ctg	cgc	tat	tgc	tat	cac	gtt	gca	ggc	gtt	gtc	ggc	ttg	atg	3197				
Asp	Thr	Leu	Arg	Tyr	Cys	Tyr	His	Val	Ala	Gly	Val	Val	Gly	Leu	Met					
	990					995					1000									
atg	gcg	caa	atc	atg	ggc	gtg	cgg	gat	aac	gcc	acg	ctg	gac	cgc		3242				
Met	Ala	Gln	Ile	Met	Gly	Val	Arg	Asp	Asn	Ala	Thr	Leu	Asp	Arg						
					1010					1015										
gcc	tgt	gac	ctt	ggg	ctg	gca	ttt	cag	ttg	acc	aat	att	gct	cgc		3287				
Ala	Cys	Asp	Leu	Gly	Leu	Ala	Phe	Gln	Leu	Thr	Asn	Ile	Ala	Arg						
	1020				1025					1030										
gat	att	gtg	gac	gat	gcg	cat	gcg	ggc	cgc	tgt	tat	ctg	ccg	gca		3332				
Asp	Ile	Val	Asp	Asp	Ala	His	Ala	Gly	Arg	Cys	Tyr	Leu	Pro	Ala						



Nostoc.ST25.txt  
1045

1035	1040		3377
agc tgg ctg gag cat gaa ggt ctg aac aaa gag aat tat gcg gca			
Ser Trp Leu Glu His Glu Gly Leu Asn Lys Glu Asn Tyr Ala Ala			
1050	1055	1060	
cct gaa aac cgt cag gcg ctg agc cgt atc gcc cgt cgt ttg gtg			3422
Pro Glu Asn Arg Gln Ala Leu Ser Arg Ile Ala Arg Arg Leu Val			
1065	1070	1075	
cag gaa gca gaa cct tac tat ttg tct gcc aca gcc ggc ctg gca			3467
Gln Glu Ala Glu Pro Tyr Tyr Leu Ser Ala Thr Ala Gly Leu Ala			
1080	1085	1090	
ggg ttg ccc ctg cgt tcc gcc tgg gca atc gct acg gcg aag cag			3512
Gly Leu Pro Leu Arg Ser Ala Trp Ala Ile Ala Thr Ala Lys Gln			
1095	1100	1105	
gtt tac cgg aaa ata ggt gtc aaa gtt gaa cag gcc ggt cag caa			3557
Val Tyr Arg Lys Ile Gly Val Lys Val Glu Gln Ala Gly Gln Gln			
1110	1115	1120	
tgg gat cag cgg cag tca acg acc acg ccc gaa aaa tta acg			3602
Trp Asp Gln Arg Gln Ser Thr Thr Thr Pro Glu Lys Leu Thr			
1125	1130	1135	
ctg ctg ctg gcc gcc tct ggt cag gcc ctt act tcc cgg atg cgg			3647
Leu Leu Leu Ala Ala Ser Gly Gln Ala Leu Thr Ser Arg Met Arg			
1140	1145	1150	
gct cat cct ccc cgc cct gcg cat ctc tgg cag cgc ccg ctc			3689
Ala His Pro Pro Arg Pro Ala His Leu Trp Gln Arg Pro Leu			
1155	1160	1165	
tagcgccatg tctttcccg agcgtcgct gaagttttga caggggcggc gcatagagga			3749
agccaaaaga aacacaacct tctttgcccc tgacggcgtg atgcatacgg tgcgccatat			3809
acaaccgttt gaggtagccc ttgcgtggaa tatagcggaa tggccaacgt tgatgcacca			3869
gcccgtcgtg caccataaaa tagagtaatc catacgccgt catacctgcg ccaatccact			3929
ggagcggcca cattcctgta ctgcccagat aaatcagcag gatcgataat gcagcaaaaa			3989
ccggcata aagatcgta acttcaaacg cacctttacg cggttcatga tgtgaaagat			4049
atccccca accccagccg tgcattgatgt atttgtgtgc cagtgcagca atcacttcca			4109
tgccaatcac ggtaacgaaa acgatcaggg cattccaaat ccacaacata atttctccgg			4169
tagagacgtc tggcagcagg cttaaggatt caattttaac agagattagc cgatctggcg			4229
gcgggaaggg aaaaaggcgc gccagaaagg cgcgccaggg atcagaagtc ggctttcaga			4289
accacacggt agttggcttt acctgcacga acatggtcca gtgcatcggt gattttcgac			4349
atcggaaggt actccactgt cggcgcaata tctgtacggc cagccagctt cagcagtga			4409
cgcagctgcg caggtgaacc ggttgaagaa cccgtcacgg cgcggtcgcc taaaatcagg			4469
ctgaaagccg ggcacgtcaa acggcttcag tacggcaccc acggtatgga acttaccg			4529
aggcgccagg gccgcaaagt agggttgcca gtcgagatcg acggcgaccg tgctgataat			4589
caggtcaaac tggcccgcga ggctttttta agctt			4624

Nostoc.ST25.txt

<211> 380

<212> PRT

<213> Erwinia uredovora

<400> 25

Met Gln Pro His Tyr Asp Leu Ile Leu Val Gly Ala Gly Leu Ala Asn  
1 5 10 15

Gly Leu Ile Ala Leu Arg Leu Gln Gln Gln Gln Pro Asp Met Arg Ile  
20 25 30

Leu Leu Ile Asp Ala Ala Pro Gln Ala Gly Gly Asn His Thr Trp Ser  
35 40 45

His His Asp Asp Leu Thr Glu Ser Gln His Arg Trp Ile Ala Pro  
50 55 60

Leu Val Val His His Trp Pro Asp Tyr Gln Val Arg Phe Pro Thr Arg  
65 70 75 80

Arg Arg Lys Leu Asn Ser Gly Tyr Phe Cys Ile Thr Ser Gln Arg Phe  
85 90 95

Ala Glu Val Leu Gln Arg Gln Phe Gly Pro His Leu Trp Met Asp Thr  
100 105 110

Ala Val Ala Glu Val Asn Ala Glu Ser Val Arg Leu Lys Lys Gly Gln  
115 120 125

Val Ile Gly Ala Arg Ala Val Ile Asp Gly Arg Gly Tyr Ala Ala Asn  
130 135 140

Ser Ala Leu Ser Val Gly Phe Gln Ala Phe Ile Gly Gln Glu Trp Arg  
145 150 155 160

Leu Ser His Pro His Gly Leu Ser Ser Pro Ile Ile Met Asp Ala Thr  
165 170 175

Val Asp Gln Gln Asn Gly Tyr Arg Phe Val Tyr Ser Leu Pro Leu Ser  
180 185 190

Pro Thr Arg Leu Leu Ile Glu Asp Thr His Tyr Ile Asp Asn Ala Thr  
195 200 205

Leu Asp Pro Glu Cys Ala Arg Gln Asn Ile Cys Asp Tyr Ala Ala Gln  
210 215 220

Gln Gly Trp Gln Leu Gln Thr Leu Leu Arg Glu Glu Gln Gly Ala Leu  
Seite 34

Nostoc.ST25.txt  
235

225

230

240

Pro Ile Thr Leu Ser Gly Asn Ala Asp Ala Phe Trp Gln Gln Arg Pro  
245 250 255

Leu Ala Cys Ser Gly Leu Arg Ala Gly Leu Phe His Pro Thr Thr Gly  
260 265 270

Tyr Ser Leu Pro Leu Ala Val Ala Val Ala Asp Arg Leu Ser Ala Leu  
275 280 285

Asp Val Phe Thr Ser Ala Ser Ile His His Ala Ile Thr His Phe Ala  
290 295 300

Arg Glu Arg Trp Gln Gln Gln Gly Phe Phe Arg Met Leu Asn Arg Met  
305 310 315 320

Phe Leu Ala Gly Pro Ala Asp Ser Arg Trp Arg Val Met Gln Arg  
325 330 335

Phe Tyr Gly Leu Pro Glu Asp Leu Ile Ala Arg Phe Tyr Ala Gly Lys  
340 345 350

Leu Thr Leu Thr Asp Arg Leu Arg Ile Leu Ser Gly Lys Pro Pro Val  
355 360 365

Pro Val Leu Ala Ala Leu Gln Ala Ile Met Thr Thr  
370 375 380

<210> 26

<211> 492

<212> PRT

<213> Erwinia uredovora

<400> 26

Met Lys Pro Thr Thr Val Ile Gly Ala Gly Phe Gly Gly Leu Ala Leu  
1 5 10 15

Ala Ile Arg Leu Gln Ala Ala Gly Ile Pro Val Leu Leu Leu Glu Gln  
20 25 30

Arg Asp Lys Pro Gly Gly Arg Ala Tyr Val Tyr Glu Asp Gln Gly Phe  
35 40 45

Thr Phe Asp Ala Gly Pro Thr Val Ile Thr Asp Pro Ser Ala Ile Glu  
50 55 60

Glu Leu Phe Ala Leu Ala Gly Lys Gln Leu Lys Glu Tyr Val Glu Leu  
Seite 35

65

70

80

Leu Pro Val Thr Pro Phe Tyr Arg Leu Cys Trp Glu Ser Gly Lys Val  
85 90 95

Phe Asn Tyr Asp Asn Asp Gln Thr Arg Leu Glu Ala Gln Ile Gln Gln  
100 105 110

Phe Asn Pro Arg Asp Val Glu Gly Tyr Arg Gln Phe Leu Asp Tyr Ser  
115 120 125

Arg Ala Val Phe Lys Glu Gly Tyr Leu Lys Leu Gly Thr Val Pro Phe  
130 135 140

Leu Ser Phe Arg Asp Met Leu Arg Ala Ala Pro Gln Leu Ala Lys Leu  
145 150 155 160

Ala Trp Arg Ser Val Tyr Ser Lys Val Ala Ser Tyr Ile Glu Asp  
165 170 175

Glu His Leu Arg Gln Ala Phe Ser Phe His Ser Leu Leu Val Gly Gly  
180 185 190

Asn Pro Phe Ala Thr Ser Ser Ile Tyr Thr Leu Ile His Ala Leu Glu  
195 200 205

Arg Glu Trp Gly Val Trp Phe Pro Arg Gly Gly Thr Gly Ala Leu Val  
210 215 220

Gln Gly Met Ile Lys Leu Phe Gln Asp Leu Gly Gly Glu Val Val Leu  
225 230 235 240

Asn Ala Arg Val Ser His Met Glu Thr Thr Gly Asn Lys Ile Glu Ala  
245 250 255

Val His Leu Glu Asp Gly Arg Arg Phe Leu Thr Gln Ala Val Ala Ser  
260 265 270

Asn Ala Asp Val Val His Thr Tyr Arg Asp Leu Leu Ser Gln His Pro  
275 280 285

Ala Ala Val Lys Gln Ser Asn Lys Leu Gln Thr Lys Arg Met Ser Asn  
290 295 300

Ser Leu Phe Val Leu Tyr Phe Gly Leu Asn His His His Asp Gln Leu  
305 310 315 320

Ala His His Thr Val Cys Phe Gly Pro Arg Tyr Arg Glu Leu Ile Asp  
325 330 335

Glu Ile Phe Asn His Asp Gly Leu Ala Glu Asp Phe Ser Leu Tyr Leu  
Seite 36

340

345

350

His Ala Pro Cys Val Thr Asp Ser Ser Leu Ala Pro Glu Gly Cys Gly  
 355 360 365

Ser Tyr Tyr Val Leu Ala Pro Val Pro His Leu Gly Thr Ala Asn Leu  
 370 375 380

Asp Trp Thr Val Glu Gly Pro Lys Leu Arg Asp Arg Ile Phe Ala Tyr  
 385 390 395 400

Leu Glu Gln His Tyr Met Pro Gly Leu Arg Ser Gln Leu Val Thr His  
 405 410 415

Arg Met Phe Thr Pro Phe Asp Phe Arg Asp Gln Leu Asn Ala Tyr His  
 420 425 430

Ser Ala Phe Ser Val Glu Pro Val Leu Thr Gln Ser Ala Trp Phe  
 435 440 445

Arg Pro His Asn Arg Asp Lys Thr Ile Thr Asn Leu Tyr Leu Val Gly  
 450 455 460

Ala Gly Thr His Pro Gly Ala Gly Ile Pro Gly Val Ile Gly Ser Ala  
 465 470 475 480

Lys Ala Thr Ala Gly Leu Met Leu Glu Asp Leu Ile  
 485 490

&lt;210&gt; 27

&lt;211&gt; 296

&lt;212&gt; PRT

&lt;213&gt; Erwinia uredovora

&lt;400&gt; 27

Met Ala Val Gly Ser Lys Ser Phe Ala Thr Ala Ser Lys Leu Phe Asp  
 1 5 10 15

Ala Lys Thr Arg Arg Ser Val Leu Met Leu Tyr Ala Trp Cys Arg His  
 20 25 30

Cys Asp Asp Val Ile Asp Asp Gln Thr Leu Gly Phe Gln Ala Arg Gln  
 35 40 45

Pro Ala Leu Gln Thr Pro Glu Gln Arg Leu Met Gln Leu Glu Met Lys  
 50 55 60

Thr Arg Gln Ala Tyr Ala Gly Ser Gln Met His Glu Pro Ala Phe Ala  
 Seite 37

65

70

80

Ala Phe Gln Glu Val Ala Met Ala His Asp Ile Ala Pro Ala Tyr Ala  
85 90 95

Phe Asp His Leu Glu Gly Phe Ala Met Asp Val Arg Glu Ala Gln Tyr  
100 105 110

Ser Gln Leu Asp Asp Thr Leu Arg Tyr Cys Tyr His Val Ala Gly Val  
115 120 125

Val Gly Leu Met Met Ala Gln Ile Met Gly Val Arg Asp Asn Ala Thr  
130 135 140

Leu Asp Arg Ala Cys Asp Leu Gly Leu Ala Phe Gln Leu Thr Asn Ile  
145 150 155 160

Arg Asp Ile Val Asp Asp Ala His Ala Gly Arg Cys Tyr Leu Pro  
165 170 175

Ala Ser Trp Leu Glu His Glu Gly Leu Asn Lys Glu Asn Tyr Ala Ala  
180 185 190

Pro Glu Asn Arg Gln Ala Leu Ser Arg Ile Ala Arg Arg Leu Val Gln  
195 200 205

Glu Ala Glu Pro Tyr Tyr Leu Ser Ala Thr Ala Gly Leu Ala Gly Leu  
210 215 220

Pro Leu Arg Ser Ala Trp Ala Ile Ala Thr Ala Lys Gln Val Tyr Arg  
225 230 235 240

Lys Ile Gly Val Lys Val Glu Gln Ala Gly Gln Gln Ala Trp Asp Gln  
245 250 255

Arg Gln Ser Thr Thr Thr Pro Glu Lys Leu Thr Leu Leu Leu Ala Ala  
260 265 270

Ser Gly Gln Ala Leu Thr Ser Arg Met Arg Ala His Pro Pro Arg Pro  
275 280 285

Ala His Leu Trp Gln Arg Pro Leu  
290 295

<210> 28

<211> 32

<212> DNA

<213> Künstliche Sequenz

Nostoc.ST25.txt

<220>

<221> primer\_bind

<222> (1)..(32)

<223>

<400> 28  
tttttctcga gcgataaacg ctcaacttggt ta

32

<210> 29

<211> 32

<212> DNA

<213> künstliche Sequenz

<220>

<221> primer\_bind

<222> (1)..(32)

<223>

<400> 29  
tttttgtcga cacgttatgc tcacaacccc gg

32

<210> 30

<211> 679

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (87)..(635)

<223>

<400> 30  
ctcgagcgat aaacgctcac ttgggttaatc atttcactct tcaattatct ataattgatga

60

gtgatcagaa ttacatgtga gaaatt atg caa acg gaa cac gtc att tta ttg  
Met Gln Thr Glu His Val Ile Leu Leu  
1 5

113

aat gca cag gga gtt ccc acg ggt acg ctg gaa aag tat gcc gca cac  
Asn Ala Gln Gly Val Pro Thr Gly Thr Leu Glu Lys Tyr Ala Ala His  
seite 39

161

Nostoc.ST25.txt  
20

10	15	20	25	
acg gca gac acc cgc tta cat ctc gcg ttc tcc agt tgg ctg ttt aat				209
Thr Ala Asp Thr Arg Leu His Leu Ala Phe Ser Ser Trp Leu Phe Asn				
	30	35	40	
gcc aaa gga caa tta tta gtt acc cgc cgc gca ctg agc aaa aaa gca				257
Ala Lys Gly Gln Leu Leu Val Thr Arg Arg Ala Leu Ser Lys Lys Ala				
	45	50	55	
tgg cct ggc gtg tgg act aac tcg gtt tgt ggg cac cca caa ctg gga				305
Trp Pro Gly Val Trp Thr Asn Ser Val Cys Gly His Pro Gln Leu Gly				
	60	65	70	
gaa agc aac gaa gac gca gtg atc cgc cgt tgc cgt tat gag ctt ggc				353
Glu Ser Asn Glu Asp Ala Val Ile Arg Arg Cys Arg Tyr Glu Leu Gly				
	75	80	85	
gtg gaa att acg cct cct gaa tct atc tat cct gac ttt cgc tac cgc				401
Val Glu Ile Thr Pro Pro Glu Ser Ile Tyr Pro Asp Phe Arg Tyr Arg				
	90	95	100	
acc gat ccg agt ggc att gtg gaa aat gaa gtg tgt ccg gta ttt				449
Thr Asp Pro Ser Gly Ile Val Glu Asn Glu Val Cys Pro Val Phe				
	110	115	120	
gcc gca cgc acc act agt gcg tta cag atc aat gat gat gaa gtg atg				497
Ala Ala Arg Thr Ser Ala Leu Gln Ile Asn Asp Asp Glu Val Met				
	125	130	135	
gat tat caa tgg tgt gat tta gca gat gta tta cac ggt att gat gcc				545
Asp Tyr Gln Trp Cys Asp Leu Ala Asp Val Leu His Gly Ile Asp Ala				
	140	145	150	
acg ccg tgg gcg ttc agt ccg tgg atg gtg atg cag gcg aca aat cgc				593
Thr Pro Trp Ala Phe Ser Pro Trp Met Val Met Gln Ala Thr Asn Arg				
	155	160	165	
gaa gcc aga aaa cga tta tct gca ttt acc cag ctt aaa taa				635
Glu Ala Arg Lys Arg Leu Ser Ala Phe Thr Gln Leu Lys				
	170	175	180	
aaaaaccccg acatttgccg gggttgtgag cataacgtgt cgac				679

> 31  
 <211> 182  
 <212> PRT  
 <213> Escherichia coli

<400> 31  
 Met Gln Thr Glu His Val Ile Leu Leu Asn Ala Gln Gly Val Pro Thr  
 1 5 10 15  
 Gly Thr Leu Glu Lys Tyr Ala Ala His Thr Ala Asp Thr Arg Leu His  
 20 25 30  
 Leu Ala Phe Ser Ser Trp Leu Phe Asn Ala Lys Gly Gln Leu Leu Val  
 35 40 45



Nostoc.ST25.txt

Thr Arg Arg Ala Leu Ser Lys Lys Ala Trp Pro Gly Val Trp Thr Asn  
 50 55 60

Ser Val Cys Gly His Pro Gln Leu Gly Glu Ser Asn Glu Asp Ala Val  
 65 70 75 80

Ile Arg Arg Cys Arg Tyr Glu Leu Gly Val Glu Ile Thr Pro Pro Glu  
 85 90 95

Ser Ile Tyr Pro Asp Phe Arg Tyr Arg Ala Thr Asp Pro Ser Gly Ile  
 100 105 110

Val Glu Asn Glu Val Cys Pro Val Phe Ala Ala Arg Thr Thr Ser Ala  
 115 120 125

Gln Ile Asn Asp Asp Glu Val Met Asp Tyr Gln Trp Cys Asp Leu  
 130 135 140

Ala Asp Val Leu His Gly Ile Asp Ala Thr Pro Trp Ala Phe Ser Pro  
 145 150 155 160

Trp Met Val Met Gln Ala Thr Asn Arg Glu Ala Arg Lys Arg Leu Ser  
 165 170 175

Ala Phe Thr Gln Leu Lys  
 180

<210> 32

<211> 31

<212> DNA

<213> Künstliche Sequenz

<220>

<221> primer\_bind

<222> (1)..(31)

<223>

<400> 32  
 tttttccatg gtgaaggagg aaatagcgaa a

31

<210> 33

<211> 32

<212> DNA

<213> Künstliche Sequenz

Nostoc.ST25.txt

<220>

<221> primer\_bind

<222> (1)..(32)

<223>

<400> 33  
ttttaagct ttcacttttt tcttgtaacc aa

32

<210> 34

<211> 962

<212> DNA

<213> Archaeoglobus fulgidus

<220>

<221> CDS

<222> (3)..(956)

<223>

<400> 34  
cc atg gtg aag gag gaa ata gcg aaa agg gcc gaa ata atc aac aaa  
Met Val Lys Glu Glu Ile Ala Lys Arg Ala Glu Ile Ile Asn Lys  
1 5 10 15

47

gcc att gaa gag ctt ctg ccc gaa agg gag ccg att gga ctc tac aaa  
Ala Ile Glu Glu Leu Leu Pro Glu Arg Glu Pro Ile Gly Leu Tyr Lys  
20 25 30

95

gac gca agg cat ctg atc aaa gca ggt ggc aag agg cta agg cct gta  
Ala Ala Arg His Leu Ile Lys Ala Gly Gly Lys Arg Leu Arg Pro Val  
35 40 45

143

ata agc ctc tta gca gtc gaa gcc ctt ggg aaa gac tac aga aag att  
Ile Ser Leu Leu Ala Val Glu Ala Leu Gly Lys Asp Tyr Arg Lys Ile  
50 55 60

191

atc ccg gct gct gtc agc att gaa aca atc cac aac ttc acc ctc gtg  
Ile Pro Ala Ala Val Ser Ile Glu Thr Ile His Asn Phe Thr Leu Val  
65 70 75

239

cat gac gac ata atg gac agg gac gag atg agg agg gga gtt ccg acg  
His Asp Asp Ile Met Asp Arg Asp Glu Met Arg Arg Gly Val Pro Thr  
80 85 90 95

287

gta cac agg gtt tat ggg gaa gcg acg gcc att tta gca ggc gac aca  
Val His Arg Val Tyr Gly Glu Ala Thr Ala Ile Leu Ala Gly Asp Thr  
100 105 110

335

ctc ttt gct gaa gcc ttc aag ctg ctg aca aag tgc gat gtt gag agc  
Leu Phe Ala Glu Ala Phe Lys Leu Leu Thr Lys Cys Asp Val Glu Ser

383

Nostoc.ST25.txt

115	120	125	
gag gga atc aga aaa gct aca gaa atg ctt tcg gac gtt tgc ata aaa Glu Gly Ile Arg Lys Ala Thr Glu Met Leu Ser Asp Val Cys Ile Lys 130 135 140			431
ata tgc gag ggg cag tac tac gac atg agc ttt gag aaa aag gag agc Ile Cys Glu Gly Gln Tyr Tyr Asp Met Ser Phe Glu Lys Lys Glu Ser 145 150 155			479
gtt tcc gag gag gag tat ctc agg atg gtc gag ctg aag acc gga gtg Val Ser Glu Glu Glu Tyr Leu Arg Met Val Glu Leu Lys Thr Gly Val 160 165 170 175			527
ctg att gca gct tct gca gca tta cct gcg gtg ctt ttt ggg gag agc Leu Ile Ala Ala Ser Ala Ala Leu Pro Ala Val Leu Phe Gly Glu Ser 180 185 190			575
gag gaa att gta aag gcg ctg tgg gac tac gga gtt ctt agc ggt att Glu Glu Ile Val Lys Ala Leu Trp Asp Tyr Gly Val Leu Ser Gly Ile 195 200 205			623
ttc cag atc cag gac gac ctg ctt gac ctg act gag gag acc gga Phe Gln Ile Gln Asp Asp Leu Leu Asp Leu Thr Glu Glu Thr Gly 210 215 220			671
aag gac tgg gga agc gac ctg ctt aaa ggg aag aaa acc ctg att gtc Lys Asp Trp Gly Ser Asp Leu Leu Lys Gly Lys Lys Thr Leu Ile Val 225 230 235			719
ata aag gcg ttc gaa aag gga gtg aag cta aag acg ttt gga aag gaa Ile Lys Ala Phe Glu Lys Gly Val Lys Leu Lys Thr Phe Gly Lys Glu 240 245 250 255			767
aag gcg gac gtc tct gag att aga gat gat atc gaa aag tta aga gag Lys Ala Asp Val Ser Glu Ile Arg Asp Asp Ile Glu Lys Leu Arg Glu 260 265 270			815
tgt ggt gcg att gat tac gct gcc agc atg gca aga aag atg gct gaa Cys Gly Ala Ile Asp Tyr Ala Ala Ser Met Ala Arg Lys Met Ala Glu 275 280 285			863
gag gcg aaa aga aag ctc gaa gtt ctg cct gaa agc aaa gcc aag gaa Glu Ala Lys Arg Lys Leu Glu Val Leu Pro Glu Ser Lys Ala Lys Glu 290 295 300			911
ca ctg ctg gaa ctt acc gac ttc ttg gtt aca aga aaa aag tga Thr Leu Leu Glu Leu Thr Asp Phe Leu Val Thr Arg Lys Lys 305 310 315			956
aagctt			962

<210> 35

<211> 317.

<212> PRT

<213> Archaeoglobus fulgidus

<400> 35

Met Val Lys Glu Glu Ile Ala Lys Arg Ala Glu Ile Ile Asn Lys Ala  
1 5 10 15

Nostoc.ST25.txt

Ile Glu Glu Leu Leu Pro Glu Arg Glu Pro Ile Gly Leu Tyr Lys Ala  
20 25 30

Ala Arg His Leu Ile Lys Ala Gly Gly Lys Arg Leu Arg Pro Val Ile  
35 40 45

Ser Leu Leu Ala Val Glu Ala Leu Gly Lys Asp Tyr Arg Lys Ile Ile  
50 55 60

Pro Ala Ala Val Ser Ile Glu Thr Ile His Asn Phe Thr Leu Val His  
65 70 75 80

Asp Asp Ile Met Asp Arg Asp Glu Met Arg Arg Gly Val Pro Thr Val  
85 90 95

Arg Val Tyr Gly Glu Ala Thr Ala Ile Leu Ala Gly Asp Thr Leu  
100 105 110

Phe Ala Glu Ala Phe Lys Leu Leu Thr Lys Cys Asp Val Glu Ser Glu  
115 120 125

Gly Ile Arg Lys Ala Thr Glu Met Leu Ser Asp Val Cys Ile Lys Ile  
130 135 140

Cys Glu Gly Gln Tyr Tyr Asp Met Ser Phe Glu Lys Lys Glu Ser Val  
145 150 155 160

Ser Glu Glu Glu Tyr Leu Arg Met Val Glu Leu Lys Thr Gly Val Leu  
165 170 175

Ile Ala Ala Ser Ala Ala Leu Pro Ala Val Leu Phe Gly Glu Ser Glu  
180 185 190

Ile Val Lys Ala Leu Trp Asp Tyr Gly Val Leu Ser Gly Ile Gly  
195 200 205

Phe Gln Ile Gln Asp Asp Leu Leu Asp Leu Thr Glu Glu Thr Gly Lys  
210 215 220

Asp Trp Gly Ser Asp Leu Leu Lys Gly Lys Lys Thr Leu Ile Val Ile  
225 230 235 240

Lys Ala Phe Glu Lys Gly Val Lys Leu Lys Thr Phe Gly Lys Glu Lys  
245 250 255

Ala Asp Val Ser Glu Ile Arg Asp Asp Ile Glu Lys Leu Arg Glu Cys  
260 265 270

Gly Ala Ile Asp Tyr Ala Ala Ser Met Ala Arg Lys Met Ala Glu Glu  
275 280 285

Nostoc.ST25.txt

Ala Lys Arg Lys Leu Glu Val Leu Pro Glu Ser Lys Ala Lys Glu Thr  
290 295 300

Leu Leu Glu Leu Thr Asp Phe Leu Val Thr Arg Lys Lys  
305 310 315

<210> 36

<211> 1293

<212> DNA

<213> Archaeoglobus fulgidus

<220>

<221> CDS

<222> (206)..(1159)

<223>

<400> 36		
taaaacgacg gccagtgagc gcgcgtaata cgactcacta tagggcgaat tgggtaccgg		60
gccccccctc gacgccgtcg ttcaatgaga atggataaga ggctcgtggg attgacgtga		120
gggggacagg atggctatat ttctgggagc gaactccggg cgaggatcta gttgtaggga		180
gggattcatg acaccacaaa cagcc atg gtg aag gag gaa ata gcg aaa agg		232
	Met Val Lys Glu Glu Ile Ala Lys Arg	
	1 5	
gcc gaa ata atc aac aaa gcc att gaa gag ctt ctg ccc gaa agg gag		280
Ala Glu Ile Ile Asn Lys Ala Ile Glu Glu Leu Pro Glu Arg Glu		
10 15 20 25		
att gga ctc tac aaa gcc gca agg cat ctg atc aaa gca ggt ggc		328
Ile Gly Leu Tyr Lys Ala Ala Arg His Leu Ile Lys Ala Gly Gly		
	30 35 40	
aag agg cta agg cct gta ata agc ctc tta gca gtc gaa gcc ctt ggg		376
Lys Arg Leu Arg Pro Val Ile Ser Leu Leu Ala Val Glu Ala Leu Gly		
	45 50 55	
aaa gac tac aga aag att atc ccg gct gct gtc agc att gaa aca atc		424
Lys Asp Tyr Arg Lys Ile Ile Pro Ala Ala Val Ser Ile Glu Thr Ile		
	60 65 70	
cac aac ttc acc ctc gtg cat gac gac ata atg gac agg gac gag atg		472
His Asn Phe Thr Leu Val His Asp Asp Ile Met Asp Arg Asp Glu Met		
	75 80 85	
agg agg gga gtt ccg acg gta cac agg gtt tat ggg gaa gcg acg gcc		520
Arg Arg Gly Val Pro Thr Val His Arg Val Tyr Gly Glu Ala Thr Ala		
	90 95 100 105	
att tta gca ggc gac aca ctc ttt gct gaa gcc ttc aag ctg ctg aca		568
Ile Leu Ala Gly Asp Thr Leu Phe Ala Glu Ala Phe Lys Leu Leu Thr		
	110 115 120	

Nostoc.ST25.txt

aag tgc gat gtt gag agc gag gga atc aga aaa gct aca gaa atg ctt Lys Cys Asp Val Glu Ser Glu Gly Ile Arg Lys Ala Thr Glu Met Leu 125 130 135	616
tcg gac gtt tgc ata aaa ata tgc gag ggg cag tac tac gac atg agc Ser Asp Val Cys Ile Lys Ile Cys Glu Gly Gln Tyr Tyr Asp Met Ser 140 145 150	664
ttt gag aaa aag gag agc gtt tcc gag gag gag tat ctc agg atg gtc Phe Glu Lys Lys Glu Ser Val Ser Glu Glu Tyr Leu Arg Met Val 155 160 165	712
gag ctg aag acc gga gtg ctg att gca gct tct gca gca tta cct gcg Glu Leu Lys Thr Gly Val Leu Ile Ala Ala Ser 180 Ala Ala Leu Pro Ala 170 175 185	760
gtg ctt ttt ggg gag agc gag gaa att gta aag gcg ctg tgg gac tac Val Leu Phe Gly Glu Ser Glu Glu Ile Val Lys Ala Leu Trp Asp Tyr 190 195 200	808
gga gtt ctt agc ggt att ggc ttc cag atc cag gac gac ctg ctt gac Val Leu Ser Gly Ile Gly Phe Gln Ile Gln Asp Asp Leu Leu Asp 205 210 215	856
ctg act gag gag acc gga aag gac tgg gga agc gac ctg ctt aaa ggg Leu Thr Glu Glu Thr Gly Lys Asp Trp Gly Ser Asp Leu Leu Lys Gly 220 225 230	904
aag aaa acc ctg att gtc ata aag gcg ttc gaa aag gga gtg aag cta Lys Lys Thr Leu Ile Val Ile Lys Ala Phe Glu Lys Gly Val Lys Leu 235 240 245	952
aag acg ttt gga aag gaa aag gcg gac gtc tct gag att aga gat gat Lys Thr Phe Gly Lys Glu Lys Ala Asp Val Ser Glu Ile Arg Asp Asp 250 255 260 265	1000
atc gaa aag tta aga gag tgt ggt gcg att gat tac gct gcc agc atg Ile Glu Lys Leu Arg Glu Cys Gly Ala Ile Asp Tyr Ala Ala Ser Met 270 275 280	1048
gca aga aag atg gct gaa gag gcg aaa aga aag ctc gaa gtt ctg cct Ala Arg Lys Met Ala Glu Glu Ala Lys Arg Lys Leu Glu Val Leu Pro 285 290 295	1096
agc aaa gcc aag gaa aca ctg ctg gaa ctt acc gac ttc ttg gtt Ser Lys Ala Lys Glu Thr Leu Leu Glu Leu Thr Asp Phe Leu Val 300 305 310	1144
aca aga aaa aag tga aagcttcaat tgcattgctct agatgatcaa agaattcctg Thr Arg Lys Lys 315	1199
gcctagtcta taggagggttt tgaaaagaaa ggagcaataa tcattttctt gttctatcaa	1259
gagggtgcta ttgctccttt ctttttttct cgag	1293

<210> 37

<211> 317

<212> PRT

<213> Archaeoglobus fulgidus

<400> 37

Nostoc.ST25.txt

Met Val Lys Glu Glu Ile Ala Lys Arg Ala Glu Ile Ile Asn Lys Ala  
1 5 10 15

Ile Glu Glu Leu Leu Pro Glu Arg Glu Pro Ile Gly Leu Tyr Lys Ala  
20 25 30

Ala Arg His Leu Ile Lys Ala Gly Gly Lys Arg Leu Arg Pro Val Ile  
35 40 45

Ser Leu Leu Ala Val Glu Ala Leu Gly Lys Asp Tyr Arg Lys Ile Ile  
50 55 60

Pro Ala Ala Val Ser Ile Glu Thr Ile His Asn Phe Thr Leu Val His  
65 70 75 80

Asp Ile Met Asp Arg Asp Glu Met Arg Arg Gly Val Pro Thr Val  
85 90 95

His Arg Val Tyr Gly Glu Ala Thr Ala Ile Leu Ala Gly Asp Thr Leu  
100 105 110

Phe Ala Glu Ala Phe Lys Leu Leu Thr Lys Cys Asp Val Glu Ser Glu  
115 120 125

Gly Ile Arg Lys Ala Thr Glu Met Leu Ser Asp Val Cys Ile Lys Ile  
130 135 140

Cys Glu Gly Gln Tyr Tyr Asp Met Ser Phe Glu Lys Lys Glu Ser Val  
145 150 155 160

Ser Glu Glu Glu Tyr Leu Arg Met Val Glu Leu Lys Thr Gly Val Leu  
165 170 175

Ala Ala Ser Ala Ala Leu Pro Ala Val Leu Phe Gly Glu Ser Glu  
180 185 190

Glu Ile Val Lys Ala Leu Trp Asp Tyr Gly Val Leu Ser Gly Ile Gly  
195 200 205

Phe Gln Ile Gln Asp Asp Leu Leu Asp Leu Thr Glu Glu Thr Gly Lys  
210 215 220

Asp Trp Gly Ser Asp Leu Leu Lys Gly Lys Lys Thr Leu Ile Val Ile  
225 230 235 240

Lys Ala Phe Glu Lys Gly Val Lys Leu Lys Thr Phe Gly Lys Glu Lys  
245 250 255

Ala Asp Val Ser Glu Ile Arg Asp Asp Ile Glu Lys Leu Arg Glu Cys  
260 265 270

Nostoc.ST25.txt

Gly Ala Ile Asp Tyr Ala Ala Ser Met Ala Arg Lys Met Ala Glu Glu  
275 280 285

Ala Lys Arg Lys Leu Glu Val Leu Pro Glu Ser Lys Ala Lys Glu Thr  
290 295 300

Leu Leu Glu Leu Thr Asp Phe Leu Val Thr Arg Lys Lys  
305 310 315

<210> 38

<211> 35

<212> DNA

<213> künstliche Sequenz

<220>

<221> primer\_bind

<222> (1)..(35)

<223>

<400> 38  
gagctcttca ttatttcgat ttgatttcg tgacc

35

<210> 39

<211> 44

<212> DNA

<213> künstliche Sequenz

<220>

<221> primer\_bind

<222> (1)..(44)

<223>

<400> 39  
aagcttgagc tcggttgatc agaagaagaa gaagaagatg aact

44

<210> 40

<211> 653

<212> DNA

<213> Arabidopsis thaliana



Nostoc.ST25.txt

<220>

<221> promoter

<222> (1)..(653)

<223>

```

<400> 40
gagctcttca ttatttcgat ttgatttcg tgaccagcga acgcagaata cttgttgtg 60
taatacttta cccgtgtaaa tcaaaaacaa aaaggctttt gagctttttg tagttgaatt 120
tctctggctg atcttttctg tacagattca tataatctga gagacgatat cattgattat 180
ttgagcttct tttgaactat ttcgtgtaat ttgggatgag agctctatgt atgtgtgtaa 240
ttgaaga caacaagaaa ggtaacaagt gagggagggg tgactccatg tcaaaataga 300
tgcataaga ggcccatcaa taagtgcctg agcccattag ctagcccagt aactaccaga 360
ttgtgagatg gatgtgtgaa cagttttttt ttgatgtag gactgaaatg tgaacaacag 420
gcgcatgaaa ggctaaatta ggacaatgat aagcagaaat aacttatcct ctctaact 480
tggcctcaca ttgcccttca cacaatccac acacatccaa tcacaacctc atcatatatc 540
tcccgctaata ctttttttct ttgatctttt tttttttgct tattattttt ttgactttga 600
tctcccatca gttcatcttc ttcttcttct tctgatcaac cgagctcaag ctt 653

```

<210> 41

<211> 28

<212> DNA

<213> Künstliche Sequenz

<220>

<221> primer\_bind

<222> (1)..(28)

<223>

```

<400> 41
gagctcactc actgatttcc attgcttg

```

28

<210> 42

<211> 30

<212> DNA

<213> Künstliche Sequenz

Nostoc.ST25.txt

<220>

<221> primer\_bind

<222> (1)..(30)

<223>

<400> 42  
aagcttgagc tctttgttga agagatttgg

30

<210> 43

<211> 37

<212> DNA

<213> Künstliche Sequenz

<220>

<221> primer\_bind

<222> (1)..(37)

<223>

<400> 43  
cgccgttaag tcgatgtccg ttgatttaaa cagtgtc

37

<210> 44

<211> 34

<212> DNA

<213> Künstliche Sequenz

<220>

<221> primer\_bind

<222> (1)..(34)

<223>

<400> 44  
atcaacggac atcgacttaa cggcgtttgt aaac

34

<210> 45

<211> 783

Nostoc.ST25.txt

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> promoter

<222> (1)..(783)

<223>

<400> 45		
gagctcactc actgatttcc attgcttgaa aattgatgat gaactaagat caatccatgt		60
tagtttcaaa acaacagtaa ctgtggccaa cttagttttg aaacaacact aactgggtcga		120
aaaaga aaaaagagtt tcatcatata tctgatttga tggactgttt ggagtttagga		180
ccaaacatta tctacaaaca aagacttttc tcctaacttg tgattccttc ttaaacccta		240
ggggtaatat tctattttcc aaggatcttt agttaaaggc aaatccggga aattattgta		300
atcatttggg gaaacatata aaagatttga gttagatgga agtgacgatt aatccaaaca		360
tatatatctc tttctttctta tttcccaaata taacagacaa aagtagaata ttggctttta		420
acaccaatat aaaaacttgc ttcacaccta aacacttttg tttacttttag ggtaagtgca		480
aaaagccaac caaatccacc tgcactgatt tgacgtttac aaacgccgtt aagtcgatgt		540
ccgttgattt aaacagtgtc ttgtaattaa aaaaatcagt ttacataaat ggaaaattta		600
tcacttagtt ttcatcaact tctgaactta cttttcatgg attaggcaat actttccatt		660
tttagtaact caagtggacc ctttacttct tcaactccat ctctctcttt ctatttcact		720
tctttcttct cattatatct cttgtcctct ccaccaaata tcttcaaca agagctcaag		780
		783

<210> 46

<211> 804

<212> DNA

<213> Synechococcus WH8102

<220>

<221> CDS

<222> (1)..(804)

<223>

<400> 46

Nostoc.ST25.txt

atg Met 1	aaa Lys	acg Thr	aca Thr	aga Arg 5	tct Ser	att Ile	tcg Ser	tgg Trp	cca Pro 10	tcg Ser	act Thr	tgc Cys	tgg Trp	cat His 15	cac His	48
cag Gln	ccg Pro	agt Ser	tgc Cys 20	tca Ser	agc Ser	tgg Trp	gtg Val	gca Ala 25	aat Asn	gag Glu	ttc Phe	agc Ser	cct Pro 30	cag Gln	gcc Ala	96
ctc Leu	aaa Lys	ggg Gly 35	ttg Leu	gct Ala	ctg Leu	gct Ala	ggg Gly 40	ctg Leu	att Ile	gga Gly	tca Ser	gcc Ala 45	tgg Trp	ctg Leu	ctc Leu	144
tcc Ser	ctg Leu 50	ggc Gly	ctg Leu	agc Ser	tac Tyr	acc Thr 55	ctg Leu	cca Pro	ctt Leu	gat Asp	cag Gln 60	acg Thr	cct Pro	ggg Gly	ctg Leu	192
ttg Leu 65	att Ile	ggc Gly	agc Ser	ttg Leu	att Ile 70	ctg Leu	ctc Leu	aga Arg	gca Ala	ttt Phe 75	ctg Leu	cac His	acc Thr	ggg Gly	ctg Leu 80	240
ttc Ile	atc Val	gtt Ala	gcc Ala	cac His 85	gat Asp	tcc Ser	atg Met	cac His	gcc Ala 90	agt Ser	ctg Leu	gtt Val	ccg Pro	ggg Gly 95	cat His	288
ccc Pro	gga Gly	ttg Leu	aac Asn 100	cgc Arg	tgg Trp	atc Ile	ggc Gly	aaa Lys 105	gtg Val	tat Tyr	ttg Leu	ttg Leu	gtg Val 110	tat Tyr	gca Ala	336
ggc Gly	ttg Leu	tct Ser 115	tat Tyr	gag Glu	cgt Arg	tgt Cys	tcc Ser 120	cgc Arg	aac Asn	cac His	aga Arg	cgt Arg 125	cat His	cac His	ctg Leu	384
gca Ala 130	ccg Pro	gag Glu	acg Thr	ttc Phe	cag Gln	gat Asp 135	cct Pro	gac Asp	tac Tyr	caa Gln	cgt Arg 140	tgc Cys	acc Thr	aat Asn	aac Asn	432
aac Asn 145	atc Ile	cta Leu	gat Asp	tgg Trp	tat Tyr 150	gtt Val	cac His	ttc Phe	atg Met	ggc Gly 155	aac Asn	tat Tyr	ctg Leu	ggc Gly	atg Met 160	480
cgg Arg	caa Gln	ctg Leu	tta Leu	aat Asn 165	cta Leu	agc Ser	tgt Cys	ctt Leu	tgg Trp 170	ctg Leu	gcg Ala	cta Leu	atc Ile	att Ile 175	ctc Leu	528
ggg Gly	tct Ser	gat Asp 180	ctc Leu	cct Pro	gct Ala	cag Gln	atc Ile 185	atg Met	cat His	ctg Leu	ctg Leu	ttg Leu 190	ttc Phe	agc Ser		576
gtt Val	ctg Leu	ccg Pro 195	ttg Leu	atc Ile	atc Ile	agt Ser	tcc Ser 200	tgt Cys	caa Gln	ttg Leu	ttt Phe	cta Leu 205	gtg Val	gga Gly	acc Thr	624
tgg Trp 210	tta Leu	ccc Pro	cac His	cga Arg	cgt Arg	ggg Gly 215	gcc Ala	acg Thr	aca Thr	cga Arg	ccg Pro 220	ggc Gly	gtg Val	aca Thr	acg Thr	672
cgc Arg 225	agc Ser	ctg Leu	gct Ala	ttg Leu	cat His 230	cca Pro	gcc Ala	ctc Leu	tct Ser	ttc Phe 235	gca Ala	gct Ala	tgt Cys	tac Tyr	aac Asn 240	720
ttt Phe	ggc Gly	tat Tyr	cat His	cgt Arg 245	gaa Glu	cat His	cat His	gaa Glu	tgc Ser 250	cct Pro	tcc Ser	aca Thr	ccc Pro	tgg Trp 255	ttt Phe	768
cag Gln	ctg Leu	cca Pro	caa Gln 260	ctt Leu	cga Arg	aat Asn	gaa Glu	tca Ser 265	ttc Phe	act Thr	tga					804

Nostoc.ST25.txt

<210> 47

<211> 267

<212> PRT

<213> Synechococcus WH8102

<400> 47

Met Lys Thr Thr Arg Ser Ile Ser Trp Pro Ser Thr Cys Trp His His  
1 5 10 15

Gln Pro Ser Cys Ser Ser Trp Val Ala Asn Glu Phe Ser Pro Gln Ala  
20 25 30

Lys Gly Leu Ala Leu Ala Gly Leu Ile Gly Ser Ala Trp Leu Leu  
35 40 45

Ser Leu Gly Leu Ser Tyr Thr Leu Pro Leu Asp Gln Thr Pro Gly Leu  
50 55 60

Leu Ile Gly Ser Leu Ile Leu Leu Arg Ala Phe Leu His Thr Gly Leu  
65 70 75 80

Phe Ile Val Ala His Asp Ser Met His Ala Ser Leu Val Pro Gly His  
85 90 95

Pro Gly Leu Asn Arg Trp Ile Gly Lys Val Tyr Leu Leu Val Tyr Ala  
100 105 110

Gly Leu Ser Tyr Glu Arg Cys Ser Arg Asn His Arg Arg His His Leu  
115 120 125

Pro Glu Thr Phe Gln Asp Pro Asp Tyr Gln Arg Cys Thr Asn Asn  
130 135 140

Asn Ile Leu Asp Trp Tyr Val His Phe Met Gly Asn Tyr Leu Gly Met  
145 150 155 160

Arg Gln Leu Leu Asn Leu Ser Cys Leu Trp Leu Ala Leu Ile Ile Leu  
165 170 175

Asn Gly Ser Asp Leu Pro Ala Gln Ile Met His Leu Leu Leu Phe Ser  
180 185 190

Val Leu Pro Leu Ile Ile Ser Ser Cys Gln Leu Phe Leu Val Gly Thr  
195 200 205

Trp Leu Pro His Arg Arg Gly Ala Thr Thr Arg Pro Gly Val Thr Thr  
210 215 220

# Nostoc.ST25.txt

Arg Ser Leu Ala Leu His Pro Ala Leu Ser Phe Ala Ala Cys Tyr Asn  
225 230 235 240

Phe Gly Tyr His Arg Glu His His Glu Ser Pro Ser Thr Pro Trp Phe  
245 250 255

Gln Leu Pro Gln Leu Arg Asn Glu Ser Phe Thr  
260 265

<210> 48

<211> 804

<212> DNA

<213> Künstliche Variante

<220>

<221> CDS

<222> (1)..(804)

<223>

<400> 48	
atg aaa acg aca aga tct att tcg tgg cca tcg act tgc tgg cat cac	48
Met Lys Thr Thr Arg Ser Ile Ser Trp Pro Ser Thr Cys Trp His His	
1 5 10 15	
cag ccg agt tgc tca agc tgg gtg gca aat gag ttc agc cct cag gcc	96
Gln Pro Ser Cys Ser Ser Trp Val Ala Asn Glu Phe Ser Pro Gln Ala	
20 25 30	
ctc aaa ggg ttg gct ctg gct ggt ctg att gga tca gcc tgg ctg ctc	144
Leu Lys Gly Leu Ala Leu Ala Gly Leu Ile Gly Ser Ala Trp Leu Leu	
35 40 45	
tcc ctg ggc ctg agc tac acc ctg cca ctt gat cag acg cct ggg ctg	192
Ser Leu Gly Leu Ser Tyr Thr Leu Pro Leu Asp Gln Thr Pro Gly Leu	
50 55 60	
ttg att ggc agc ttg att ctg tgg cag acc ttt ctg cac acc ggg ctg	240
Leu Ile Gly Ser Leu Ile Leu Trp Gln Thr Phe Leu His Thr Gly Leu	
65 70 75 80	
ttc atc gtt gcc cac gat tcc atg cac gcc agt ctg gtt ccg ggt cat	288
Phe Ile Val Ala His Asp Ser Met His Ala Ser Leu Val Pro Gly His	
85 90 95	
ccc gga ttg aac cgc tgg atc ggc aaa gtg tat ttg ttg gtg tat gca	336
Pro Gly Leu Asn Arg Trp Ile Gly Lys Val Tyr Leu Leu Val Tyr Ala	
100 105 110	
ggc ttg tct tat gag cgt tgt tcc cgc aac cac aga cgt cat cac ctg	384
Gly Leu Ser Tyr Glu Arg Cys Ser Arg Asn His Arg Arg His His Leu	
115 120 125	
gca ccg gag acg ttc cag gat cct gac tac caa cgt tgc acc aat aac	432
Ala Pro Glu Thr Phe Gln Asp Pro Asp Tyr Gln Arg Cys Thr Asn Asn	

Nostoc.ST25.txt  
140

130 135 480  
aac atc cta gat tgg tat gtt cac ttc atg ggc aac tat ctg ggc atg  
Asn Ile Leu Asp Trp Tyr Val His Phe Met Gly Asn Tyr Leu Gly Met  
145 150 155 160

528  
cgg caa ctg tta aat cta agc tgt ctt tgg ctg gcg cta atc att ctc  
Arg Gln Leu Leu Asn Leu Ser Cys Leu Trp Leu Ala Leu Ile Ile Leu  
165 170 175

576  
aac ggt tct gat ctc cct gct cag atc atg cat ctg ctg ttg ttc agc  
Asn Gly Ser Asp Leu Pro Ala Gln Ile Met His Leu Leu Leu Phe Ser  
180 185 190

624  
gtt ctg ccg ttg atc atc agt tcc tgt caa ttg ttt cta gtg gga acc  
Val Leu Pro Leu Ile Ile Ser Ser Cys Gln Leu Phe Leu Val Gly Thr  
195 200 205

672  
tgg tta ccc cac cga cgt ggg gcc acg aca cga ccg ggc gtg aca acg  
Trp Leu Pro His Arg Arg Gly Ala Thr Thr Arg Pro Gly Val Thr Thr  
210 215 220

720  
agc ctg gct ttg cat cca gcc ctc tct ttc gca gct tgt tac aac  
Ser Leu Ala Leu His Pro Ala Leu Ser Phe Ala Ala Cys Tyr Asn  
225 230 235 240

768  
ttt ggc tat cat cgt gaa cat cat gaa tcg cct tcc aca ccc tgg ttt  
Phe Gly Tyr His Arg Glu His His Glu Ser Pro Ser Thr Pro Trp Phe  
245 250 255

804  
cag ctg cca caa ctt cga aat gaa tca ttc act tga  
Gln Leu Pro Gln Leu Arg Asn Glu Ser Phe Thr  
260 265

<210> 49

<211> 267

<212> PRT

<213> Künstliche Variante

<214> 49

Met Lys Thr Thr Arg Ser Ile Ser Trp Pro Ser Thr Cys Trp His His  
1 5 10 15

Gln Pro Ser Cys Ser Ser Trp Val Ala Asn Glu Phe Ser Pro Gln Ala  
20 25 30

Leu Lys Gly Leu Ala Leu Ala Gly Leu Ile Gly Ser Ala Trp Leu Leu  
35 40 45

Ser Leu Gly Leu Ser Tyr Thr Leu Pro Leu Asp Gln Thr Pro Gly Leu  
50 55 60

Leu Ile Gly Ser Leu Ile Leu Trp Gln Thr Phe Leu His Thr Gly Leu  
65 70 75 80

Phe Ile Val Ala His Asp Ser Met His Ala Ser Leu Val Pro Gly His  
Seite 55

Pro Gly Leu Asn Arg Trp Ile Gly Lys Val Tyr Leu Leu Val Tyr Ala  
100 105 110

Gly Leu Ser Tyr Glu Arg Cys Ser Arg Asn His Arg Arg His His Leu  
115 120 125

Ala Pro Glu Thr Phe Gln Asp Pro Asp Tyr Gln Arg Cys Thr Asn Asn  
130 135 140

Asn Ile Leu Asp Trp Tyr Val His Phe Met Gly Asn Tyr Leu Gly Met  
145 150 155 160

Arg Gln Leu Leu Asn Leu Ser Cys Leu Trp Leu Ala Leu Ile Ile Leu  
165 170 175

Gly Ser Asp Leu Pro Ala Gln Ile Met His Leu Leu Leu Phe Ser  
180 185 190

Val Leu Pro Leu Ile Ile Ser Ser Cys Gln Leu Phe Leu Val Gly Thr  
195 200 205

Trp Leu Pro His Arg Arg Gly Ala Thr Thr Arg Pro Gly Val Thr Thr  
210 215 220

Arg Ser Leu Ala Leu His Pro Ala Leu Ser Phe Ala Ala Cys Tyr Asn  
225 230 235 240

Phe Gly Tyr His Arg Glu His His Glu Ser Pro Ser Thr Pro Trp Phe  
245 250 255

Gln Leu Pro Gln Leu Arg Asn Glu Ser Phe Thr  
260 265

<210> 50

<211> 804

<212> DNA

<213> Künstliche Variante

<220>

<221> CDS

<222> (1)..(804)

<223>

<400> 50



Nostoc.ST25.txt

atg Met 1	aaa Lys	acg Thr	aca Thr	aga Arg 5	tct Ser	att Ile	tcg Ser	tgg Trp	cca Pro 10	tcg Ser	act Thr	tgc Cys	tgg Trp	cat His 15	cac His	48
cag Gln	ccg Pro	agt Ser	tgc Cys 20	tca Ser	agc Ser	tgg Trp	gtg Val	gca Ala 25	aat Asn	gag Glu	ttc Phe	agc Ser	cct Pro 30	cag Gln	gcc Ala	96
ctc Leu	aaa Lys	ggg Gly 35	ttg Leu	gct Ala	ctg Leu	gct Ala	ggt Gly 40	ctg Leu	att Ile	gga Gly	tca Ser	gcc Ala 45	tgg Trp	ctg Leu	ctc Leu	144
tcc Ser	ctg Leu 50	ggc Gly	ctg Leu	agc Ser	tac Tyr	acc Thr 55	ctg Leu	cca Pro	ctt Leu	gat Asp	cag Gln 60	acg Thr	cct Pro	ggg Gly	ctg Leu	192
ttg Leu 65	att Ile	ggc Gly	agc Ser	ttg Leu	att Ile 70	ctg Leu	ctc Leu	aga Arg	gca Ala	ttt Phe 75	ctg Leu	cac His	acc Thr	ggg Gly	ctg Leu 80	240
ttc Ile	atc Val	gtt Ala	gcc Ala	cac His 85	gat Asp	tcc Ser	atg Met	cac His	gcc Ala 90	agt Ser	ctg Leu	gtt Val	ccg Pro	ggt Gly 95	cat His	288
ccc Pro	gga Gly	ttg Leu	aac Asn 100	cgc Arg	tgg Trp	atc Ile	ggc Gly	aaa Lys 105	gtg Val	tat Tyr	ttg Leu	ttg Leu	gtg Val 110	tat Tyr	gca Ala	336
ggc Gly	ttg Leu	tct Ser 115	tat Tyr	gag Glu	cgt Arg	tgt Cys	tcc Ser 120	cgc Arg	aac Asn	cac His	aga Arg	cgt Arg 125	cat His	cac His	gga Gly	384
cat His 130	cct Pro	ggt Gly	act Thr	gat Asp	tta Leu	gat Asp 135	cct Pro	gac Asp	tac Tyr	caa Gln	cgt Arg 140	tgc Cys	acc Thr	aat Asn	aac Asn	432
aac Asn 145	atc Ile	cta Leu	gat Asp	tgg Trp	tat Tyr 150	gtt Val	cac His	ttc Phe	atg Met	ggc Gly 155	aac Asn	tat Tyr	ctg Leu	ggc Gly	atg Met 160	480
cgg Arg	caa Gln	ctg Leu	tta Leu	aat Asn 165	cta Leu	agc Ser	tgt Cys	ctt Leu	tgg Trp 170	ctg Leu	gcg Ala	cta Leu	atc Ile	att Ile 175	ctc Leu	528
ggt Gly	tct Ser	gat Asp 180	ctc Leu	cct Pro	gct Ala	cag Gln	atc Ile 185	atg Met	cat His	ctg Leu	ctg Leu	ttg Leu 190	ttc Phe	agc Ser		576
gtt Val	ctg Leu	ccg Pro 195	ttg Leu	atc Ile	atc Ile	agt Ser	tcc Ser 200	tgt Cys	caa Gln	ttg Leu	ttt Phe	cta Leu 205	gtg Val	gga Gly	acc Thr	624
tgg Trp 210	tta Leu	ccc Pro	cac His	cga Arg	cgt Arg	ggg Gly 215	gcc Ala	acg Thr	aca Thr	cga Arg	ccg Pro 220	ggc Gly	gtg Val	aca Thr	acg Thr	672
cgc Arg 225	agc Ser	ctg Leu	gct Ala	ttg Leu	cat His 230	cca Pro	gcc Ala	ctc Leu	tct Ser	ttc Phe 235	gca Ala	gct Ala	tgt Cys	tac Tyr	aac Asn 240	720
ttt Phe	ggc Gly	tat Tyr	cat His	cgt Arg 245	gaa Glu	cat His	cat His	gaa Glu	tcg Ser 250	cct Pro	tcc Ser	aca Thr	ccc Pro	tgg Trp 255	ttt Phe	768
cag Gln	ctg Leu	cca Pro	caa Gln 260	ctt Leu	cga Arg	aat Asn	gaa Glu	tca Ser 265	ttc Phe	act Thr	tga					804

Nostoc.ST25.txt

<210> 51

<211> 267

<212> PRT

<213> Künstliche Variante

<400> 51

Met Lys Thr Thr Arg Ser Ile Ser Trp Pro Ser Thr Cys Trp His His  
1 5 10 15

Gln Pro Ser Cys Ser Ser Trp Val Ala Asn Glu Phe Ser Pro Gln Ala  
20 25 30

Lys Gly Leu Ala Leu Ala Gly Leu Ile Gly Ser Ala Trp Leu Leu  
35 40 45

Ser Leu Gly Leu Ser Tyr Thr Leu Pro Leu Asp Gln Thr Pro Gly Leu  
50 55 60

Leu Ile Gly Ser Leu Ile Leu Leu Arg Ala Phe Leu His Thr Gly Leu  
65 70 75 80

Phe Ile Val Ala His Asp Ser Met His Ala Ser Leu Val Pro Gly His  
85 90 95

Pro Gly Leu Asn Arg Trp Ile Gly Lys Val Tyr Leu Leu Val Tyr Ala  
100 105 110

Gly Leu Ser Tyr Glu Arg Cys Ser Arg Asn His Arg Arg His His Gly  
115 120 125

Pro Gly Thr Asp Leu Asp Pro Asp Tyr Gln Arg Cys Thr Asn Asn  
130 135 140

Asn Ile Leu Asp Trp Tyr Val His Phe Met Gly Asn Tyr Leu Gly Met  
145 150 155 160

Arg Gln Leu Leu Asn Leu Ser Cys Leu Trp Leu Ala Leu Ile Ile Leu  
165 170 175

Asn Gly Ser Asp Leu Pro Ala Gln Ile Met His Leu Leu Leu Phe Ser  
180 185 190

Val Leu Pro Leu Ile Ile Ser Ser Cys Gln Leu Phe Leu Val Gly Thr  
195 200 205

Trp Leu Pro His Arg Arg Gly Ala Thr Thr Arg Pro Gly Val Thr Thr  
210 215 220

Nostoc.ST25.txt

Arg Ser Leu Ala Leu His Pro Ala Leu Ser Phe Ala Ala Cys Tyr Asn  
225 230 235 240

Phe Gly Tyr His Arg Glu His His Glu Ser Pro Ser Thr Pro Trp Phe  
245 250 255

Gln Leu Pro Gln Leu Arg Asn Glu Ser Phe Thr  
260 265

Abbildung 1

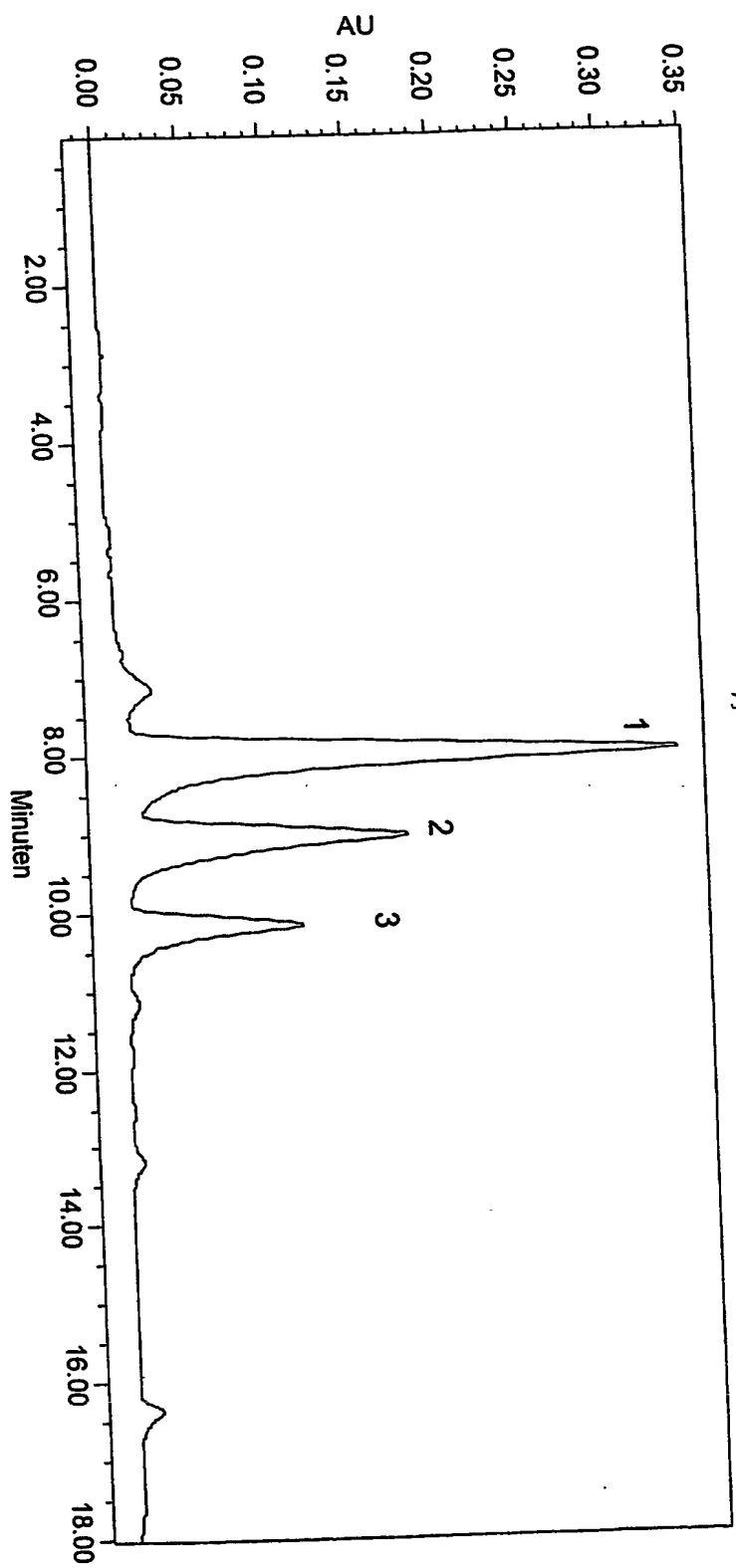


Abbildung 2

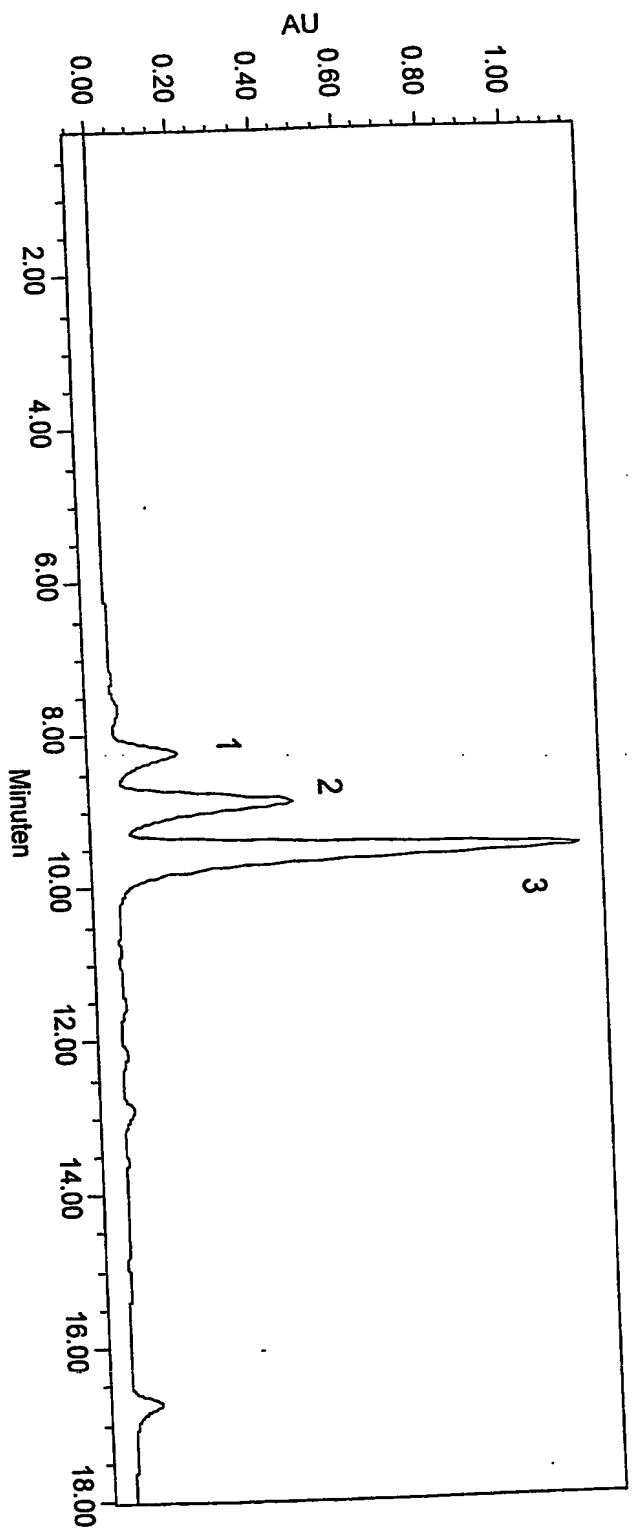


Abbildung 5

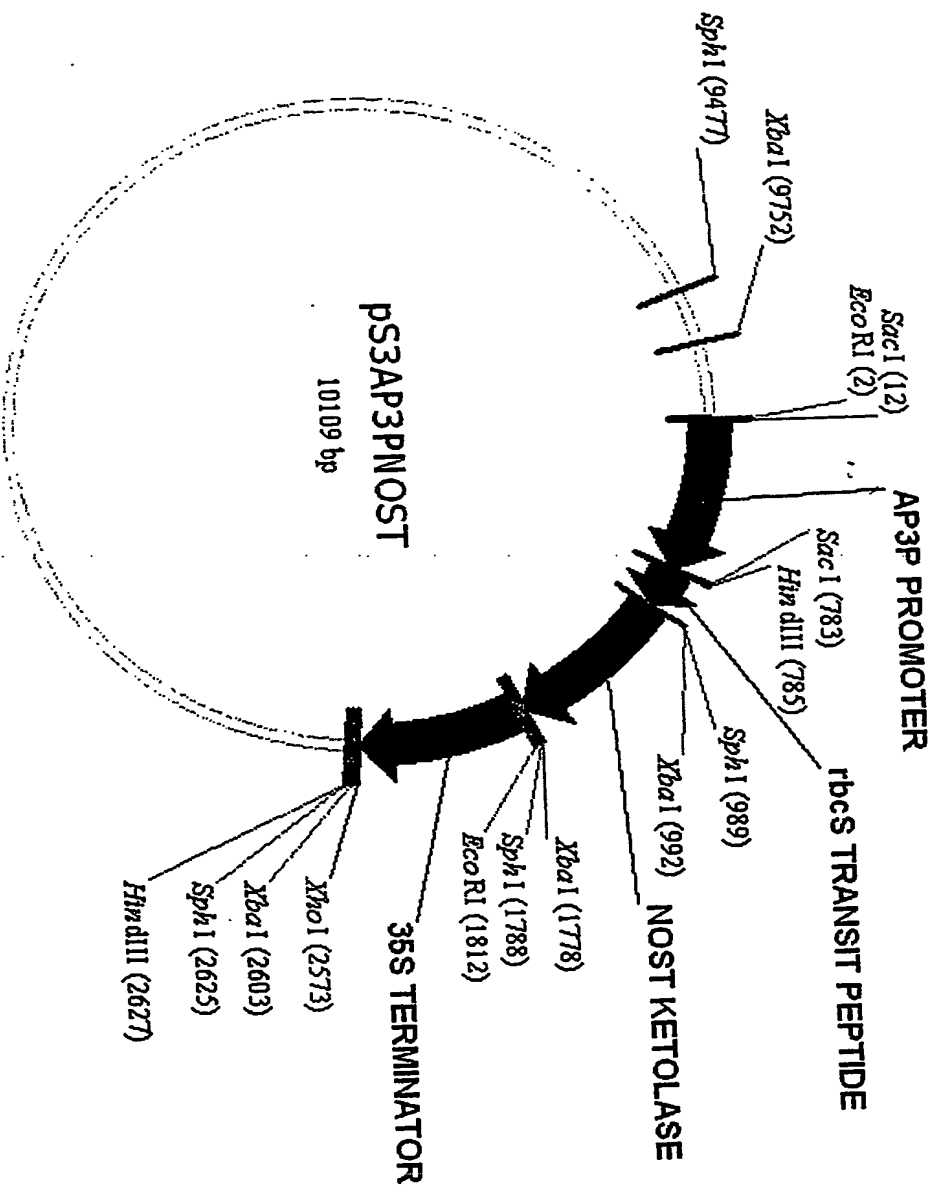
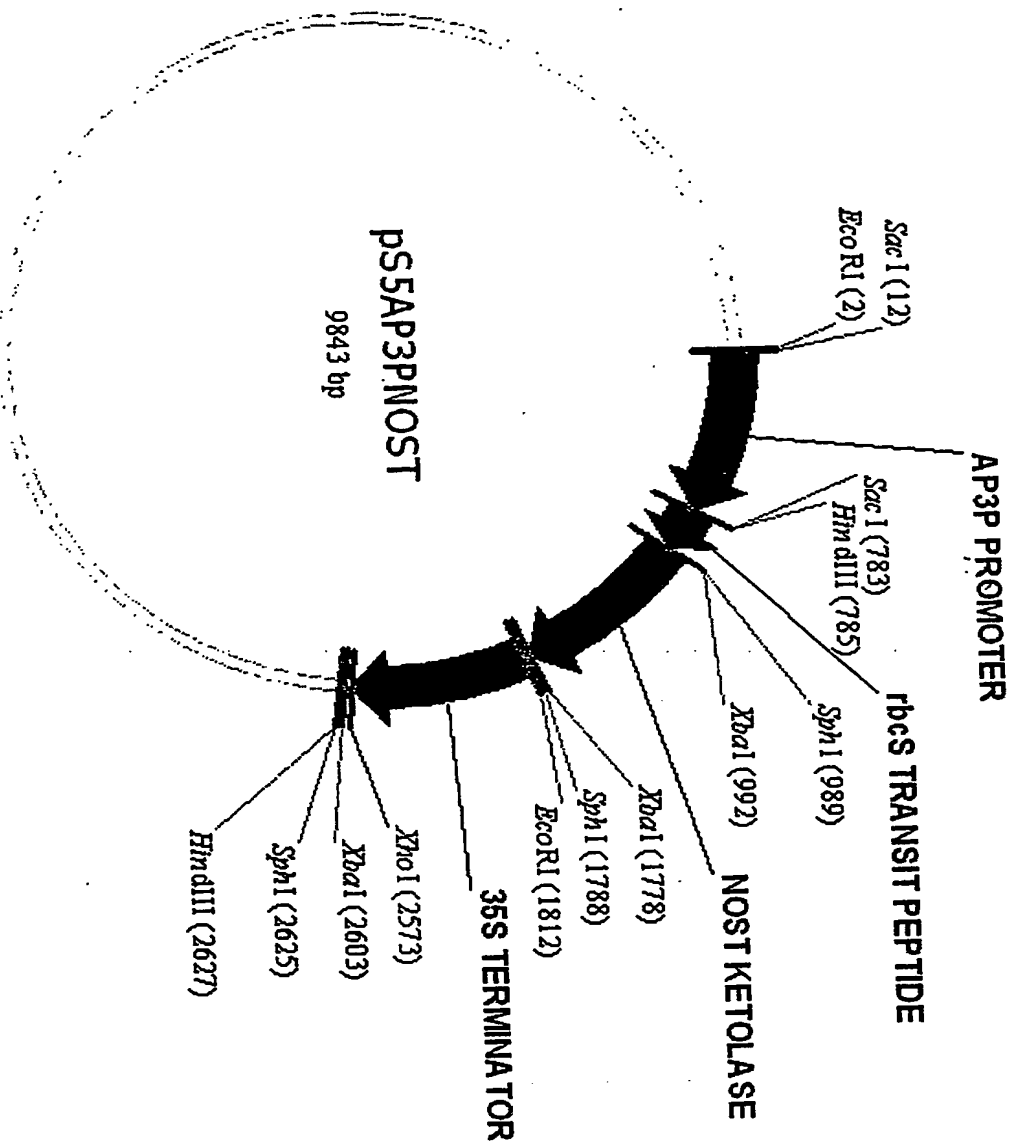


Abbildung 6



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**